



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number 114402

TO: Minh-Tam Davis
Location: REM-3C18/3A24
Art Unit: 1642
Monday, February 23, 2004

Case Serial Number: 09899569

From: Toby Port
Location: Biotech-Chem Library
Remsen 1A59
Phone: 571-272-2523

toby.port@uspto.gov

Search Notes

Dear Examiner Davis,

Here are the results of your search.
Please feel free to contact me if you have any questions.

Toby Port



This Page Blank (uspio)

XX 17-JAN-2002.
PD 05-JUL-2001; 2001WO-EP07705.
XX 07-JUL-2000; 2000DE-1031080.
XX 19-APR-2001; 2001DE-1019294.
XX (BOEH) BOEHRINGER INGELHEIM INT GMBH.
PI Schaefer N, Scherl-Mostagier M, Sommergruber W, Abseher R;
XX MPI, 2002-171704/22.
XX P-PSDB; AAM49641.
PT New tumor-associated antigen B345, useful for diagnosis and
XX immunotherapy of tumors, also related nucleic acid and antibodies -
PS Claim 3; Page 79-85; 102pp; German.
XX This invention describes a novel tumour-associated antigen, designated
CC B345 which has cytotoxic activity. B345 is involved in communication,
CC interaction and/or signal transduction with extracellular components and
CC ligands, especially in the metastatic potential of cancers, particularly
CC of the colon. B345 or its immunogenic fragments, also the DNA that
CC encodes it, are useful for immunotherapy of cancer, particularly
CC carcinoma of lung or colon. Antibodies raised against B345 are useful for
CC treatment and diagnosis of cancers that are associated with B345
CC expression, including their use for targeted delivery of cytotoxic or
CC radioactive agents. Probes derived from B345 can be used to detect
CC tumour-specific mutations in the B345 sequence, and can be used to screen
CC for B345 specific modulators. This sequence encodes a human B345
CC tumour-associated antigen described in the invention.
SQ Sequence 6163 BP; 1501 A; 1689 C; 1482 G; 1491 T; 0 other;
Query Match 100.0%; Score 6163; DB 24; Length 6163;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCACACCGCGAATGGGAGTAGTGGGACCCGACCAACCGGATGCGGGAGCCCTGACCC 60
DB 1 CCACACCGCGAATGGGAGTAGTGGGACCCGACCAACCGGATGCGGGAGCCCTGACCC 60
QY 61 TGGGAGGAGAGGCGGCTGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
DB 61 TGGGAGGAGAGGCGGCTGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 121 CGGATGAGCTCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
DB 121 CGGATGAGCTCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 121 CGGATGAGCTCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
DB 121 CGGATGAGCTCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 181 GTGAGTGAAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
DB 181 GTGAGTGAAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
QY 241 CCACCGTCTGTTTCCACCGAGGCGGAGGCTGCCGAGTCAATGAGCGGCTGAAGTGC 300
DB 241 CCACCGTCTGTTTCCACCGAGGCGGAGGCTGCCGAGTCAATGAGCGGCTGAAGTGC 300
QY 301 GGGGCTCTATCGCACTGCTAGGGGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB 301 GGGGCTCTATCGCACTGCTAGGGGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 361 GCAGAGCTTTGAGATTGCTGCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 361 GCAGAGCTTTGAGATTGCTGCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 GGGAGCCCGACTGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 421 GGGAGCCCGACTGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 481 ATGTTGTCCATCAAGTCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540

DB 481 ATGTTGTCCATCAAGTCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 541 AATCACTTGTGATGAGATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 541 AATCACTTGTGATGAGATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 601 GGGAGAGTTGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCA 660
DB 601 GGGAGAGTTGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCA 660
QY 661 TGGAGATGCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCA 720
DB 661 TGGAGATGCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCA 720
QY 721 AGGAGATGAGTCCGAGGAG 780
DB 721 AGGAGATGAGTCCGAGGAG 780
QY 781 ATGATGCCACCGTGTGTCAGATGCGAATCTTCTGAGCAATGAGCACTGTGTCCGAGATC 840
DB 781 ATGATGCCACCGTGTGTCAGATGCGAATCTTCTGAGCAATGAGCACTGTGTCCGAGATC 840
QY 841 AAGATGCCAAG 900
DB 841 AAGATGCCAAG 900
QY 901 TCCGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCA 960
DB 901 TCCGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCA 960
QY 961 TTTGAGGAGTGAAGGCTCAAGCAACCTGATGTCGCAATCAAGAGAGGCTTCCCTGAG 1020
DB 961 TTTGAGGAGTGAAGGCTCAAGCAACCTGATGTCGCAATCAAGAGAGGCTTCCCTGAG 1020
QY 1021 GATGAGCTTCATGACGTGAGAGTGTGCTGCTGCAACCTGCGGAGAGAGAGAGAGAG 1080
DB 1021 GATGAGCTTCATGACGTGAGAGTGTGCTGCTGCAACCTGCGGAGAGAGAGAGAGAG 1080
QY 1081 CTCAACTTCAACCTTCTCAACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB 1081 CTCAACTTCAACCTTCTCAACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 1141 GGGTCAACCAACCCGAGGAGTTCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
DB 1141 GGGTCAACCAACCCGAGGAGTTCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 1201 GGGAGCTTCAACCTTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB 1201 GGGAGCTTCAACCTTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY 1261 CGGCTGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAG 1320
DB 1261 CGGCTGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAG 1320
QY 1321 GTTGAATTGAGTGAAG 1380
DB 1321 GTTGAATTGAGTGAAG 1380
QY 1381 AGCCGCAAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
DB 1381 AGCCGCAAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
QY 1441 CTCAACCTGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAG 1500
DB 1441 CTCAACCTGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAG 1500
QY 1501 CTGTGATGAATGAG 1560
DB 1501 CTGTGATGAATGAG 1560
QY 1561 TCTTACTCACTCAGAGTCCAGTGAATCTTCACTGCTGCTGAGCTGATGACTTTC 1620

Db	1561	TCCTACTACTCCAGGTGCCAGTAGACATCTCTCAACCTGTGTGGAGCTGATGACTTC	1620
Qy	1621	TCCTGGAAAGCTGTGTGTGTGCCCAAGGACAGAGCTCAGCTGTGTGTGTGCCACGACCAGAG	1680
Db	1621	TCCTGGAAAGCTGTGTGTGTGCCCAAGGACAGAGCTCAGCTGTGTGTGTGTGCCAGGCCAAGAG	1680
Qy	1681	CTGCAGACAGCATACACACGAGAGAGCCCTGTGCACACACAGCTTCAGCTTACTGTGTGGCACT	1740
Db	1681	CTGCAGACAGCATACACAGAGAGAGCCCTGTGCACACACAGCTTCAGCTTACTGTGTGGCACT	1740
Qy	1741	GGCATATCCACACAGGAGCTGTATCTTGGGGCTCTTTCGGCCGGGGAGGGCTATCAACAG	1800
Db	1741	GGCATATCCACACAGGAGCTGTATCTTGGGGCTCTTTCGGCCGGGGAGGGCTATCAACAG	1800
Qy	1801	ATCCAGGTGAAGACAGAACATCTCGGTGACCCCTTGCGACTTTGTGCCCCAGCTTCAACAA	1860
Db	1801	ATCCAGGTGAAGACAGAACATCTCGGTGACCCCTTGCGACTTTGTGCCCCAGCTTCAACAA	1860
Qy	1861	GAGGCTTCAGGCAAGGCTCTGACGGTGTCTTTATACCTTATTTCAAGAGAGGGCTT	1920
Db	1861	GAGGCTTCAGGCAAGGCTCTGACGGTGTCTTTATACCTTATTTCAAGAGAGGGCTT	1920
Qy	1921	TTTCACGGTGACCCCTGACACAAAAGAGAGGTCTAACCTGAGAGACCCCACTGGGAGACGG	1980
Db	1921	TTTCACGGTGACCCCTGACACAAAAGAGAGGTCTAACCTGAGAGACCCCACTGGGAGACGG	1980
Qy	1981	GGCCTGCCATCCCTCACCCTCTGTGTCTGTGAAACATCAGCGTGGCCAGAGACAGAGTGGCC	2040
Db	1981	GGCCTGCCATCCCTCACCCTCTGTGTCTGTGAAACATCAGCGTGGCCAGAGACAGAGTGGCC	2040
Qy	2041	TGCTGACTTTCTTTAAGAGACGGAGCGGCGTGTCTTGCGACAGACAGGCGCGCATTCATG	2100
Db	2041	TGCTGACTTTCTTTAAGAGACGGAGCGGCGTGTCTTGCGACAGAGCGCGCATTCATG	2100
Qy	2101	ATCATTCACAGAGACAGCGGACCCCGGGCTGAGAGATCTTCAACCTGTGACAGAGATGTGCTC	2160
Db	2101	ATCATTCACAGAGACAGCGGACCCCGGGCTGAGAGATCTTCAACCTGTGACAGAGATGTGCTC	2160
Qy	2161	CCCAAGCCAACTTTCACATCAGACTTCTGGGTCAACATCTCTTAACTGACAGCCCAAGC	2220
Db	2161	CCCAAGCCAACTTTCACATCAGACTTCTGGGTCAACATCTCTTAACTGACAGCCCAAGC	2220
Qy	2221	AGCGGCAAGCAGTAGACCTGTCTCTCGGTGACATTAACCCCAAGACGTGTGACTTG	2280
Db	2221	AGCGGCAAGCAGTAGACCTGTCTCTCGGTGACATTAACCCCAAGACGTGTGACTTG	2280
Qy	2281	ACTGTCACTCTCATATCGCAGCGGTGGAGAGTGGAGTCTTAATCTGTGTCTGTGCCCTGTGGGCTC	2340
Db	2281	ACTGTCACTCTCATATCGCAGCGGTGGAGAGTGGAGTCTTAATCTGTGTCTGTGCCCTGTGGGCTC	2340
Qy	2341	ATCATTTTGCTGTGTGAAGAAAGAGAAAAAGAAACAAACAGGGGCCCGCTGTGGGTATC	2400
Db	2341	ATCATTTTGCTGTGTGAAGAAAGAGAAAAAGAAACAAACAGGGGCCCGCTGTGGGTATC	2400
Qy	2401	TACAAATGGCAACATCAATATCTAGATGTCCGAGGACGCAAAAAAGTTTCAAGAAAGGCGA	2460
Db	2401	TACAAATGGCAACATCAATATCTAGATGTCCGAGGACGCAAAAAAGTTTCAAGAAAGGCGA	2460
Qy	2461	AAGGACAAATGACTCCCATGTATATGCACTATCGAGACACCATGTGTATATGGCATCTGT	2520
Db	2461	AAGGACAAATGACTCCCATGTATATGCACTATCGAGACACCATGTGTATATGGCATCTGT	2520
Qy	2521	CTACAGAGATTCACAGGGGCTCTTCTCTGTGACCAAGGTGGACACTTACCTGGCGTTCAG	2580
Db	2521	CTACAGAGATTCACAGGGGCTCTTCTCTGTGACCAAGGTGGACACTTACCTGGCGTTCAG	2580
Qy	2581	GGCAACAAATGGGGTCTGTCTCTCCCTCCCAACCAACAAATAGCTTCAGAGGGCCCAACTGCA	2640
Db	2581	GGCAACAAATGGGGTCTGTCTCTCCCTCCCAACCAACAAATAGCTTCAGAGGGCCCAACTGCA	2640
Qy	2641	AAAGTTGGCACTGAGAGACCACTTCTCTGTCTCTCTCTCTGTAGTCTGAGAGTGAACGGTAC	2700
Db	2641	AAAGTTGGCACTGAGAGACCACTTCTCTGTCTCTCTCTCTGTAGTCTGAGAGTGAACGGTAC	2700

QY	2701	ACCTTCTCCATCCCAACAATGGGAGTGAAGCAGAGGACAGACATTCCTTA	CTG	2760			
Db	2701	ACCTTCTCCATCCCAACAATGGGAGTGAAGCAGAGGACAGACATTCCTTA	CTG	2760			
QY	2761	AAACCTCAGAGGCCATGAGCCGAGCAATTA	CTTGA	2820			
Db	2761	AAACCTCAGAGGCCATGAGCCGAGCAATTA	CTTGA	2820			
QY	2761	AAACCTCAGAGGCCATGAGCCGAGCAATTA	CTTGA	2820			
Db	2761	AAACCTCAGAGGCCATGAGCCGAGCAATTA	CTTGA	2820			
QY	2821	GTTCATTAAGCAGGCACTGAGACACCCGCGTGTCTTAAC	CAGAAATCCTAAAGAA	2880			
Db	2821	GTTCATTAAGCAGGCACTGAGACACCCGCGTGTCTTAAC	CAGAAATCCTAAAGAA	2880			
QY	2881	GAGGAATTATA	CAGAGGAA	2940			
Db	2881	GAGGAATTATA	CAGAGGAA	2940			
QY	2941	TCAGTGA	CTATCTTAAGGCAAGACATTGAAATTA	CTGAATCTGATACAGT	3000		
Db	2941	TCAGTGA	CTATCTTAAGGCAAGACATTGAAATTA	CTGAATCTGATACAGT	3000		
QY	3001	CATGACAGCTCATGTGCTCTCAACTTAAGGCTGCGGTTAG	CAGAGGCTGTAATGAGAGG	3060			
Db	3001	CATGACAGCTCATGTGCTCTCAACTTAAGGCTGCGGTTAG	CAGAGGCTGTAATGAGAGG	3060			
QY	3061	AGAGAGGCTGAGTCACTAGACATAGGGTTG	CAGCAAGCCCTGATTCAGAGTTTAAC	3120			
Db	3061	AGAGAGGCTGAGTCACTAGACATAGGGTTG	CAGCAAGCCCTGATTCAGAGTTTAAC	3120			
QY	3121	AGAGGCTTGCCCTCTTCAAGACAA	CAGTTCCAAATTC	CAGAGAGCTTACCTGAGGCTCCTTA	3180		
Db	3121	AGAGGCTTGCCCTCTTCAAGACAA	CAGTTCCAAATTC	CAGAGAGCTTACCTGAGGCTCCTTA	3180		
QY	3181	CTCTCAGCTGGGGTCCCAAGATGAAACGA	CAATGATCCTTTTATTA	TATTAATTTATTTG	3240		
Db	3181	CTCTCAGCTGGGGTCCCAAGATGAAACGA	CAATGATCCTTTTATTA	TATTAATTTATTTG	3240		
QY	3241	TGCTCTGCTGTTATTTAAGATCAAA	TGTAATCA	CACTAGCTCTTTCACCTGACTTA	3300		
Db	3241	TGCTCTGCTGTTATTTAAGATCAAA	TGTAATCA	CACTAGCTCTTTCACCTGACTTA	3300		
QY	3301	GTAATTA	CTACTTA	CTGGTTGGA	TGCTGGGTTGTA	CTTCTACGAGCTTACAT	3360
Db	3301	GTAATTA	CTACTTA	CTGGTTGGA	TGCTGGGTTGTA	CTTCTACGAGCTTACAT	3360
QY	3361	AAAGGTGGCCGTGCCCCCAGGTGGTGGAA	TTAATTA	CAATCTGTCACCA	CCAGAAAGA	3420	
Db	3361	AAAGGTGGCCGTGCCCCCAGGTGGTGGAA	TTAATTA	CAATCTGTCACCA	CCAGAAAGA	3420	
QY	3421	ATGTGTGTGTTGAGCAGCA	ATTGACATATCTGCTTGA	TAAAGAGACTTCGATTC	3480		
Db	3421	ATGTGTGTGTTGAGCAGCA	ATTGACATATCTGCTTGA	TAAAGAGACTTCGATTC	3480		
QY	3481	TAGGTGCGTTCGTGGTTATCCCAT	TTGTGGAAATTCATCTTGAATCCCAT	TGTCCTATAGT	3540		
Db	3481	TAGGTGCGTTCGTGGTTATCCCAT	TTGTGGAAATTCATCTTGAATCCCAT	TGTCCTATAGT	3540		
QY	3541	CTTGAAGCAATAGAAATTTCTCA	ATTTCCCAATGTCGGGTTCTCTTA	AGCTGAGCAATA	3600		
Db	3541	CTTGAAGCAATAGAAATTTCTCA	ATTTCCCAATGTCGGGTTCTCTTA	AGCTGAGCAATA	3600		
QY	3601	CTTTGACATTTAAAGAGAAATTTGAGAA	TAATTTCTACCTCTTA	AAAAATGTTTAAATATA	3660		
Db	3601	CTTTGACATTTAAAGAGAAATTTGAGAA	TAATTTCTACCTCTTA	AAAAATGTTTAAATATA	3660		
QY	3661	TACCAAA	CAGTGGCCCCCTGCATTA	GTGTTTCTGTGGCA	CTGCAACCCATTTA	CTTGTTAG	3720
Db	3661	TACCAAA	CAGTGGCCCCCTGCATTA	GTGTTTCTGTGGCA	CTGCAACCCATTTA	CTTGTTAG	3720
QY	3721	CTTAAAA	CAACAATTA	AGCTTA	TAAGTCC	TAAGAAATGAGATG	3780
Db	3721	CTTAAAA	CAACAATTA	AGCTTA	TAAGTCC	TAAGAAATGAGATG	3780

3781 GAATGAAATCAAGGTGTCAGCAGAGCTGTGCTCCTTCTGAAAGCTCTAGAGGAAGCCG 3840
3781 GAATGAAATCAAGGTGTCAGCAGAGCTGTGCTCCTTCTGAAAGCTCTAGAGGAAGCCG 3840
3841 GTTCCCTGCAATTTCAAGGCTTCTAGAGGCTGGCTGATTTCCAGAGCTCCAGTGGCTGCTC 3900
3841 GTTCCCTGCAATTTCAAGGCTTCTAGAGGCTGGCTGATTTCCAGAGCTCCAGTGGCTGCTC 3900
3901 AAGCTTTTCTCAGCATGTCAGCTGTCAGCTGGCCCTCCACTTCCCTTTTGAATTAC 3960
3901 AAGCTTTTCTCAGCATGTCAGCTGTCAGCTGGCCCTCCACTTCCCTTTTGAATTAC 3960
3961 AAAGCCACACAGAGAGATCCAGAGTATCTCTCCATCTAAAGATCCTTATCATCTCTGGA 4020
3961 AAAGCCACACAGAGAGATCCAGAGTATCTCTCCATCTAAAGATCCTTATCATCTCTGGA 4020
4021 AAGAGCTTTTGCATGCAAGACACATAGCCACAGGTGGGATTTAGAGCCAGAGCATCTT 4080
4021 AAGAGCTTTTGCATGCAAGACACATAGCCACAGGTGGGATTTAGAGCCAGAGCATCTT 4080
4081 TGGGGTGTGTTATTTCTGCTACACACACCTTCTGCTGCACTGCACTCCACAGAGAGGCTA 4140
4081 TGGGGTGTGTTATTTCTGCTACACACACCTTCTGCTGCACTGCACTCCACAGAGAGGCTA 4140
4141 CAAATGATCTGGCGCACAGGGATGTTTGTGTTAGCTTGGGACCTTAACTTAAAAA 4200
4141 CAAATGATCTGGCGCACAGGGATGTTTGTGTTAGCTTGGGACCTTAACTTAAAAA 4200
4201 ACCCGAGTCAGAGATCTGGGCATGCTGGGGCTCATTTCTCACTTACCAACATCTGGC 4260
4201 ACCCGAGTCAGAGATCTGGGCATGCTGGGGCTCATTTCTCACTTACCAACATCTGGC 4260
4261 TGGAGCTGGGCAACAGCTGTGCTTTAGAGAGGGGTGTCACTTCAACAGGTCAACACAGC 4320
4261 TGGAGCTGGGCAACAGCTGTGCTTTAGAGAGGGGTGTCACTTCAACAGGTCAACACAGC 4320
4321 TGGAGCTGGGCAACAGCTGTGCTTTAGAGAGGGGTGTCACTTCAACAGGTCAACACAGC 4380
4321 TGGAGCTGGGCAACAGCTGTGCTTTAGAGAGGGGTGTCACTTCAACAGGTCAACACAGC 4380
4381 CCCCTGAGGTTCATTTATGTAAGAAAGAAAGAGCTGGGATTAATCTGTAATCAAG 4440
4381 CCCCTGAGGTTCATTTATGTAAGAAAGAAAGAGCTGGGATTAATCTGTAATCAAG 4440
4441 TGAAGTACATGAGACCAATGTGTCTCACTTACCTTTTCTTTTCTTTTCTTTTCT 4500
4441 TGAAGTACATGAGACCAATGTGTCTCACTTACCTTTTCTTTTCTTTTCTTTTCT 4500
4501 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 4560
4501 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 4560
4561 TGGCGCAATCTGGGCTCACTGCAACCTCTGCTCTGGGCTCAAGGAATTTCTCCACCTC 4620
4561 TGGCGCAATCTGGGCTCACTGCAACCTCTGCTCTGGGCTCAAGGAATTTCTCCACCTC 4620
4621 AAGCTTCCCAATAGCTGGGATCTGGCAAAACCAACATGCCAGCTAATTTTGTATTT 4680
4621 AAGCTTCCCAATAGCTGGGATCTGGCAAAACCAACATGCCAGCTAATTTTGTATTT 4680
4681 TTTGTAGAGACAGGGTTTCAACATGTTGCTCAAGGCTGCTCAACCTCTGGGCTCAAGC 4740
4681 TTTGTAGAGACAGGGTTTCAACATGTTGCTCAAGGCTGCTCAACCTCTGGGCTCAAGC 4740
4741 AATTCCTGCTGGGCTCTCCCAAGTGTGGGATTTACAGATGTGACCAACGATCCAG 4800
4741 AATTCCTGCTGGGCTCTCCCAAGTGTGGGATTTACAGATGTGACCAACGATCCAG 4800
4801 CCCACACCTCATTTATACCAATTTACCTGCCAGTACCTGTGGATCTTTTGTCTCTCAG 4860
4801 CCCACACCTCATTTATACCAATTTACCTGCCAGTACCTGTGGATCTTTTGTCTCTCAG 4860
4861 CCTGCTCTGATCTGAGAGAGAGGATTAATGTTATAGCTTGTCAACACATCTCCAAAGTT 4920

4861 CCTGCTCTGATCTGAGAGAGAGGATTAATGTTATGCTGTGACACAGTCCAAAGTT 4920
4921 CAATATTTCTGGGCAAAAATTCTTCAAAAAAATAATGTACTCATTTGATTAATGA 4980
4921 CAATATTTCTGGGCAAAAATTCTTCAAAAAAATAATGTACTCATTTGATTAATGA 4980
4981 ATTCACTTTGAAATGCAACCGCTTCACTTGTTCATGAGCTAAATGAAAGAAATTTTA 5040
4981 ATTCACTTTGAAATGCAACCGCTTCACTTGTTCATGAGCTAAATGAAAGAAATTTTA 5040
5041 TAGTCTCTTAATGAGCTGTGTAAGAGACCTTTGAAACATTTTCAAGAGATAGATAT 5100
5041 TAGTCTCTTAATGAGCTGTGTAAGAGACCTTTGAAACATTTTCAAGAGATAGATAT 5100
5101 TTAAGTCATGCTGCTGGGCTTGCATGAGCACTTCTCTTCTGAAAGTCTGGTTCTGTC 5160
5101 TTAAGTCATGCTGCTGGGCTTGCATGAGCACTTCTCTTCTGAAAGTCTGGTTCTGTC 5160
5161 CCAAGTACCTTTGAGCTTGTAGAGCCGAGATGTCACCTGCTAAAGGCTCAAGAGAGG 5220
5161 CCAAGTACCTTTGAGCTTGTAGAGCCGAGATGTCACCTGCTAAAGGCTCAAGAGAGG 5220
5221 CTGGGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5280
5221 CTGGGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5280
5281 CCTCATTTCTGACATCTCCCACTCCAGGCTTCTTCAAGAGAGAGCTGAGCTGCTG 5340
5281 CCTCATTTCTGACATCTCCCACTCCAGGCTTCTTCAAGAGAGAGCTGAGCTGCTG 5340
5341 ATTCCACCAAGGTGGATTTGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5400
5341 ATTCCACCAAGGTGGATTTGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5400
5401 CACTGTGCTGCAAGAGACACAGCTGGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5460
5401 CACTGTGCTGCAAGAGACACAGCTGGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5460
5461 TGGAGAGTGTGCTGCTGCTGAGACCAATTTCTTCTGAGCTGTGCTGAGAGAGGCT 5520
5461 TGGAGAGTGTGCTGCTGCTGAGACCAATTTCTTCTGAGCTGTGCTGAGAGAGGCT 5520
5521 CCAGCTGAGGAACCTGGGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5580
5521 CCAGCTGAGGAACCTGGGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5580
5581 TCAAGGTCACTGGAAACAGTGAAGTGCATTTGTGAAGCTTCACTGACATGCCAGCT 5640
5581 TCAAGGTCACTGGAAACAGTGAAGTGCATTTGTGAAGCTTCACTGACATGCCAGCT 5640
5641 GCTCATCAAGTGTGCTGCTGAGCACTGAGCAAGGAGGCAAGGCACTGAGGATCTGCT 5700
5641 GCTCATCAAGTGTGCTGCTGAGCACTGAGCAAGGAGGCAAGGCACTGAGGATCTGCT 5700
5701 AATGCTGTGCTATTGCAACAGAGGAAAGTCTCAAGGAAGTCAACTGGGACAGCA 5760
5701 AATGCTGTGCTATTGCAACAGAGGAAAGTCTCAAGGAAGTCAACTGGGACAGCA 5760
5761 AATGCTGTGCTATTGCAACAGAGGAAAGTCTCAAGGAAGTCAACTGGGACAGCA 5820
5761 AATGCTGTGCTATTGCAACAGAGGAAAGTCTCAAGGAAGTCAACTGGGACAGCA 5820
5821 GCTTCACTGAGAAATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 5880
5821 GCTTCACTGAGAAATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 5880
5881 CTGTCTTTTAAATAAATAATGATCAAAAAAAGATCAAGCTGTTTAACTGCTCGG 5940
5881 CTGTCTTTTAAATAAATAATGATCAAAAAAAGATCAAGCTGTTTAACTGCTCGG 5940
5941 CTACTTGTCCCTGGGTTCAATGAGAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6000
5941 CTACTTGTCCCTGGGTTCAATGAGAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6000

DB 5941 CTACTGTCCTCCGTTGACAGAGAGCCCGGTTTCCAGTTGTCAGTGCAGAGCTC 6000
 QY 6001 AGCATGGCTCAGCAGATGCTGTCTTAATTGTGATGATACAGAAAGCCAGCTTTGGG 6060
 DB 6001 AGCATGGCTCAGCAGATGCTGTCTTAATTGTGATGATACAGAAAGCCAGCTTTGGG 6060
 QY 6061 ATACAGTTCTTCTCTTCATTTGATGCGGTGACCTGTGTGAGACAGATGTTTTGTCC 6120
 DB 6061 ATACAGTTCTTCTCTTCATTTGATGCGGTGACCTGTGTGAGACAGATGTTTTGTCC 6120
 QY 6121 GGAATTAATAATATAGCTTGGAGTCTGCGCCAAAAA 6163
 DB 6121 GGAATTAATAATATAGCTTGGAGTCTGCGCCAAAAA 6163

RESULT 2

ABA99506
 ID ABA99506 standard; cDNA, 5897 BP.

ABA99506;
 AC

17-MAY-2002 (first entry)

Human tumour-associated antigen B345 cDNA.

Tumour-associated antigen; human; B345; cytosolic; cell communication;
 cell interaction; signal transduction; metastasis; cancer; colon;
 immunotherapy; carcinoma; lung; diagnosis; gene; ss.

Homo sapiens.

Location/Qualifiers

1..214

5'UTR

/*tag= a

215..2464

/*tag= b

/product= "tumour-associated antigen B345"

2465..5897

/*tag= c

3'UTR

WO200204508-A1.

17-JAN-2002.

05-JUL-2001; 2001WO-BP07705.

07-JUL-2000; 2000DE-1031080.

19-APR-2001; 2001DE-1019294.

(BOEH) BOEHRINGER INGELHEIM INT GMBH.

Schweifer N, Scherl-Mostageer M, Sommergruber W, Abseher R;

WPI; 2002-171704/22.

P-PSDB; AAM49640.

New tumour-associated antigen B345, useful for diagnosis and

immunotherapy of tumors, also related nucleic acid and antibodies

Example 6; Page 70-76; 102pp; German.

This invention describes a novel tumour-associated antigen, designated B345 which has cytotoxic activity. B345 is involved in communication, interaction and/or signal transduction with extracellular components and ligands, especially in the metastatic potential of cancers, particularly of the colon. B345 or its immunogenic fragments, also the DNA that encodes it, are useful for immunotherapy of cancer, particularly carcinoma of lung or colon. Antibodies raised against B345 are useful for treatment and diagnosis of cancers that are associated with B345 expression, including their use for targeted delivery of cytotoxic or radioactive agents. Probes derived from B345 can be used to detect tumour-specific mutations in the B345 sequence, and can be used to screen for B345 specific modulators. This sequence encodes a human B345

CC tumour-associated antigen described in the invention.

XX Sequence 5897 BP; 1464 A; 1596 C; 1364 G; 1467 T; 6 other;

Query Match 87.5%; Score 5395; DB 24; Length 5897;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 5865; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 291 CCTGAACGCGGGGCTCTATGCACTGCTAGAGGTTGCTGCTGCGGGGCGGCGCT 350
 DB 25 CCTGAACGCGGGGCTCTATGCACTGCTAGAGGTTGCTGCTGCGGGGCGGCGCT 84
 QY 351 GCGCGCGCGGAGAGGCTTTGAGATTGCTCTGCGCAGAGAACCAATTACAGTTCT 410
 DB 85 GCGCGCGCGGAGAGGCTTTGAGATTGCTCTGCGCAGAGAACCAATTACAGTTCT 144
 QY 411 CATTAAGCTGGGAGACCCGACCTCTGCGCAAAACCTGTTACATGCTATTCTTAAAG 470
 DB 145 CATTAAGCTGGGAGACCCGACCTCTGCGCAAAACCTGTTACATGCTATTCTTAAAG 204
 QY 471 ACATTAACCATGTTGTCATCAAGTCTGAGAGAAATAGCTTACCTTTAGCGCGCA 530
 DB 205 ACATTAACCATGTTGTCATCAAGTCTGAGAGAAATAGCTTACCTTTAGCGCGCA 264
 QY 531 GAGTCTGAGATCATTTGTCATAGATCCAGAAAAATATGACTGATGTCAGGCC 590
 DB 265 GAGTCTGAGATCATTTGTCATAGATCCAGAAAAATATGACTGATGTCAGGCC 324
 QY 591 ATGTCCTTTGGGAGGTTGAGCTTCAAGCTTCGACATGCTGTTGCTTACCTCAACG 650
 DB 325 ATGTCCTTTGGGAGGTTGAGCTTCAAGCTTCGACATGCTGTTGCTTACCTCAACG 384
 QY 651 AACTTTCATGCGGATGTCAAAGCTCATTAAGCATCGGTTTAGAGCTGCACTTTCCAT 710
 DB 385 AACTTTCATGCGGATGTCAAAGCTCATTAAGCATCGGTTTAGAGCTGCACTTTCCAT 444
 QY 711 CCTTGCGCTGAGGAGCATGCTGCGGAGTGAAGCTGCCAGAGGATCACTCATCCAT 770
 DB 445 CCTTGCGCTGAGGAGCATGCTGCGGAGTGAAGCTGCCAGAGGATCACTCATCCAT 504
 QY 771 CAGCGGCGGAATGATGTCACCGTGTGATGAGATCGGAACTTTCTGACGATGGAATG 830
 DB 505 CAGCGGCGGAATGATGTCACCGTGTGATGAGATCGGAACTTTCTGACGATGGAATG 564
 QY 831 GTCCCGGATCAAGATGCAAGAGAGATGAAATGACCTTCAACCTCCATGTTCAACC 890
 DB 565 GTCCCGGATCAAGATGCAAGAGAGATGAAATGACCTTCAACCTCCATGTTCAACC 624
 QY 891 CAGAAATGCTCCGCTCAGCATTGCAAAACCGCTCATCTATAAAGCTGTGATCAT 950
 DB 625 CAGAAATGCTCCGCTCAGCATTGCAAAACCGCTCATCTATAAAGCTGTGATCAT 684
 QY 951 CGAGTCTGTGTTGAGGTTGAAGGCTCAGCAACCTGATGTTGCGCACTACCCAGAAAG 1010
 DB 685 CGAGTCTGTGTTGAGGTTGAAGGCTCAGCAACCTGATGTTGCGCACTACCCAGAAAG 744
 QY 1011 CTTCCTTAGAGATGAGCTCAGAGTGGCACTGTTGCTGCTGCAACCTCGGCGCAG 1070
 DB 745 CTTCCTTAGAGATGAGCTCAGAGTGGCACTGTTGCTGCTGCAACCTCGGCGCAG 804
 QY 1071 GGTCTCTTCTCTCAACTTCAACCTCTCAACCTGTGAGAGAGAGAGCGGGTTGAATA 1130
 DB 805 GGTCTCTTCTCTCAACTTCAACCTCTCAACCTGTGAGAGAGAGAGCGGGTTGAATA 864
 QY 1131 CTACATCCCGGCTCAGCAGCAACCCGAGGTTTCAAGCTGAGAGCAAGCAGCTGG 1190
 DB 865 CTACATCCCGGCTCAGCAGCAACCCGAGGTTTCAAGCTGAGAGCAAGCAGCTGG 924
 QY 1191 GAACTGCGGCGGAACTTCAACCTCTCTGCAAGGCTGTGACCAAGATGCCAAGTCC 1250
 DB 925 GAACTGCGGCGGAACTTCAACCTCTCTGCAAGGCTGTGACCAAGATGCCAAGTCC 984
 QY 1251 AGGATCTCTCGGCTGAGTTCAAAGTTTGTGTCACATCCAAATGAAGCAATTA 1310

Db 985 AGGAGTCCCTCCGCGTGCAGATTCCAAAGTTTGGTCCAAATCCACAAATGAAGCAATTA 1044
Qy 1311 AATCAACGTGTTGACTTGTAGTAATAGAGGAGCATGTCACTCAACATCGAGCCAGGCG 1370
Db 1045 AATCTACGTGTGTTGACTTGTAGTAATAGAGGAGCATGTCACTCAACATCGAGCCAGGCG 1104
Qy 1371 CGTCAACAGAGCCCGCAAGTTTGTCTGGCTGTGTCTGTGTCTAGAAATCTCGAGCTTG 1430
Db 1105 CGTCAAAACAGAGCCCGCAAGTTTGTCTGGCTGTGTCTGTGTCTAGAAATCTCGAGCTTG 1164
Qy 1431 CAGTAGCAACCTCACCCCTGACATCTGGGCTCCAAACGAAATCTCCCTTCTTGTGATGA 1490
Db 1465 CAGTAGCAACCTCACCCCTGACATCTGGCTCCAAACGAAATCTCCCTTCTTGTGATGA 1224
Qy 1491 TGTGACACGTCTGTGGATGAATGTGGAATAAACCATTAAGCTGCACAGACCAACCGGTA 1550
Db 1225 TGTGACACGTCTGTGGATGAATGTGGAATAAACCATTAAGCTGCACAGACCAACCGGTA 1284
Qy 1551 CCAAGAGAAATCTCACTCACTCCAGGTGCCAGTGCATCTTCACCTGCTGTGGAGCT 1610
Db 1285 CCAAGAGAAATCTCACTCACTCCAGGTGCCAGTGCATCTTCACCTGCTGTGGAGCT 1344
Qy 1611 GATAGACTTCTCTGTGAAGCTGTGTGTGCCCAAGGACAGGCTCAAGCTGTGTGTGTGCC 1670
Db 1345 GATAGACTTCTCTGTGAAGCTGTGTGTGCCCAAGGACAGGCTCAAGCTGTGTGTGTGCC 1404
Qy 1671 AGCCCGAAGCTGCAGCAGACATACACAGAGAAAGCCCTGCACACAGGCTTACGTAACCT 1730
Db 1405 AGCCCGAAGCTGCAGCAGACATACACAGAGAAAGCCCTGCACACAGGCTTACGTAACCT 1464
Qy 1731 CGTGGCAGTGCATACCCAGCCAGGACCTGTATCTTCGGCTCTCTTCGCGGAGAGCTC 1790
Db 1465 CGTGGCAGTGCATACCCAGCCAGGACCTGTATCTTCGGCTCTCTTCGCGGAGAGCTC 1524
Qy 1791 TATCAAGCAGATTCAGAGTGAAGCAAAATCTTCGGTGAACCTTCGACCTTTGGCCCGCAG 1850
Db 1525 TATCAAGCAGATTCAGAGTGAAGCAAAATCTTCGGTGAACCTTCGACCTTTGGCCCGCAG 1584
Qy 1851 CTTCACAAAGAGAGGCTTCAGGAGGAGGTCTGACGGGTCTCTTATATACCTTATTTCAAGA 1910
Db 1585 CTTCACAAAGAGAGGCTTCAGGAGGAGGTCTGACGGGTCTCTTATATACCTTATTTCAAGA 1644
Qy 1911 GGAAGGCGTTTTCAAGCGTACCCCTGCACACAAAGACAAAGGTCTACCTGAGAACCCCAA 1970
Db 1645 GGAAGGCGTTTTCAAGCGTACCCCTGCACACAAAGACAAAGGTCTACCTGAGAACCCCAA 1704
Qy 1971 CTGGGACCGGGGCTGCGCATTCCTCACTCTGTGTCTGGAACATAGCGTGCACAGAA 2030
Db 1705 CTGGGACCGGGGCTGCGCATTCCTCACTCTGTGTCTGGAACATAGCGTGCACAGAA 1764
Qy 2031 CAGAGTGGCTGCTGCTGATCTTTCTTAAAGAGCGGAGCGGTGTGTCGACAGCAGGGG 2090
Db 1765 CAGAGTGGCTGCTGCTGATCTTTCTTAAAGAGCGGAGCGGTGTGTCGACAGCAGGGG 1824
Qy 2091 CGCATTCATGATCATTCAGAGAGCAGCGAACCCGGGCTGAGAGATCTTACAGCTGAGCGA 2150
Db 1825 CGCATTCATGATCATTCAGAGAGCAGCGAACCCGGGCTGAGAGATCTTACAGCTGAGCGA 1884
Qy 2151 GGAATGTGCTCCCAAGAGCAAGCTTCCATCACTCAGAGCTTGTGGGTCAATCTTAACTG 2210
Db 1885 GGAATGTGCTCCCAAGAGCAAGCTTCCATCACTCAGAGCTTGTGGGTCAATCTTAACTG 1944
Qy 2211 CAGCCCAAGAGGAGCAAGCAGTAGACCTGCTCTTCTGGGTGACCTTAACCCCAAGAG 2270
Db 1945 CAGCCCAAGAGGAGCAAGCAGTAGACCTGCTCTTCTGGGTGACCTTAACCCCAAGAG 2004
Qy 2271 TGTGACCTTGAAGTGTCTCATCTCAGCGGTGTGGAGGTGAGTCTTACGTGTCTG 2330
Db 2005 TGTGACCTTGAAGTGTCTCATCTCAGCGGTGTGGAGGTGAGTCTTACGTGTCTG 2064
Qy 2331 CCTGGGCTCATCTTGTCTGTGTGAAAAAGAAAGAAAGACAAACAGGAGCCCGC 2390

Db 2065 CCTGGGCTCATCTTGTCTGTGTGAAAAAGAAAGAAAGACAAACAGGAGCCCGC 2124
Qy 2391 TGTGGATATCTAACATGAGCAATCAATCACTAGATGCGAGGAGCCCAAAAAAGTTTCA 2450
Db 2125 TGTGGATATCTAACATGAGCAATCAATCACTAGATGCGC-AGGAGCCAAAAAGTTTCA 2183
Qy 2451 GAAAGGGGAAAGCAATGACTCCCATGTGTATGACATCTGAGAGACACATAGGTATA 2510
Db 2184 GAAAGGGGAAAGCAATGACTCCCATGTGTATGACATCTGAGAGACACATAGGTATA 2243
Qy 2511 TGGGATCTGCTACAGATATCCAGGAGCTCTCTCTGACGACGAGGTGACACCTTACG 2570
Db 2244 TGGGATCTGCTACAGATATCCAGGAGCTCTCTCTGACGACGAGGTGACACCTTACG 2303
Qy 2571 GCCGTTCCAGGACACATGAGGAGTGTCTCTCTCTCCACCCACACATATGCTCCAGGC 2630
Db 2304 GCCGTTCCAGGACACATGAGGAGTGTCTCTCTCTCCACCCACACATATGCTCCAGGC 2363
Qy 2631 CCAACCTGCAAAAGTTGGCCACTGAGAGACCACTCTGCTCTCCCTCTGAGTCTGAGAG 2690
Db 2364 CCAACCTGCAAAAGTTGGCCACTGAGAGACCACTCTGCTCTCCCTCTGAGTCTGAGAG 2423
Qy 2691 TGAACCGTACACCTTCTCCCATCCCAACATGAGGAGTATGAGCAGCAAGACACAGACAT 2750
Db 2424 TGAACCGTACACCTTCTCCCATCCCAACATGAGGAGTATGAGCAGCAAGACACAGACAT 2483
Qy 2751 TCCCTTACTGAACTCTCAGAGCCCATGAGGCCACAGATTAATCTTATCTCAGAC 2810
Db 2484 TCCCTTACTGAACTCTCAGAGCCCATGAGGCCACAGATTAATCTTATCTCAGAC 2543
Qy 2811 GCTTTGCTGAGTTTATTAAGCAGGAGCACTGAGACACCCGTCGTTCTTAAACAGAA 2870
Db 2544 GCTTTGCTGAGTTTATTAAGCAGGAGCACTGAGACACCCGTCGTTCTTAAACAGAA 2603
Qy 2871 TCCCTTAAAGAGGAAATTAACAGAGAAACGACGAGGTTTCTCTGACACCGCAAC 2930
Db 2604 TCCCTTAAAGAGGAAATTAACAGAGAAACGACGAGGTTTCTCTGACACCGCAAC 2663
Qy 2931 TTCACATTGCTCAGTGACTCATTTCTTAAAGGCGCAACATTTGAAATGATTAATCCAAATC 2990
Db 2664 TTCACATTGCTCAGTGACTCATTTCTTAAAGGCGCAACATTTGAAATGATTAATCCAAATC 2723
Qy 2991 TGGATACAGTCACTGACAGTCACTGCTCTTCACTTAAAGCTGTGCGTTAGCCAGCTTG 3050
Db 2724 TGGATACAGTCACTGACAGTCACTGCTCTTCACTTAAAGCTGTGCGTTAGCCAGCTTG 2783
Qy 3051 TAAAGAGGAGAGAGGCTGCTGACCTGATGAGATGAGGTTGACGAAAGCCGATTCAG 3110
Db 2784 TAAAGAGGAGAGAGGCTGCTGATGAGATGAGGTTGACGAAAGCCGATTCAG 2843
Qy 3111 AGTGTAAACAGAGGCTTGCCCTCTTCAAGACAAAGTTCCAAATTCAGAGAGCCCTACCT 3170
Db 2844 AGTGTAAACAGAGGCTTGCCCTCTTCAAGACAAAGTTCCAAATTCAGAGAGCCCTACCT 2903
Qy 3171 GAGGTCCCTACTCTCACTGAGGCTCCCAAGATGAAGACAACTGTGCTTTTATATT 3230
Db 2904 GAGGTCCCTACTCTCACTGAGGCTCCCAAGATGAAGACAACTGTGCTTTTATATT 2963
Qy 3231 ATTTATTTGGGTGCTGCTGTTATTTAAGAGATCAAAATGATTAACCACTAGCTCTTTC 3290
Db 2964 ATTTATTTGGGTGCTGCTGTTATTTAAGAGATCAAAATGATTAACCACTAGCTCTTTC 3023
Qy 3291 ACCTGACTTAGTAATACATCACTACTACTGTTTGAATGCTCGGTTGTGACTTACTAG 3350
Db 3024 ACCTGACTTAGTAATACATCACTACTACTGTTTGAATGCTCGGTTGTGACTTACTAG 3083
Qy 3351 ACCGCTAGATAACGTGTGCTGTCCCAAGGTGTGGAAATTAATTAATCTGTCCAA 3410
Db 3084 ACCGCTAGATAACGTGTGCTGTCCCAAGGTGTGGAAATTAATTAATCTGTCCAA 3143
Qy 3411 CCAGAAAGAAATGTGTGCTTTGAGCAGCATGACAAATATCTGCTTGTATGAAGACTT 3470
Db 3144 CCAGAAAGAAATGTGTGCTTTGAGCAGCATGACAAATATCTGCTTGTATGAAGACTT 3203

QY 3471 CCTGATTCCTAGGTCGGTTCGGTTATCCCATTTGGAAATTCATCTTGAATCCATT 3530
DB 3204 CCGATTCCTAGGTCGGTTCGGTTATCCCATTTGGAAATTCATCTTGAATCCATT 3263
QY 3531 GTCCTATAGTCCTAGCAATTAAGAAATTTCTCAAGTTTCCATGTGCGGTCTCTTAGC 3590
DB 3264 GTCCTATAGTCCTAGCAATTAAGAAATTTCTCAAGTTTCCATGTGCGGTCTCTTAGC 3323
QY 3591 TGCACATATCTTGAATTAAGAAATTTAGAAATTTTCATCTTGAATTAAGAAATG 3650
DB 3324 TGCACATATCTTGAATTAAGAAATTTAGAAATTTTCATCTTGAATTAAGAAATG 3383
QY 3651 TTTAAATATATACCAAAAGTGGCCCTGCTAGTATGTTTCTGTTGCCATCAACCCAT 3710
DB 3384 TTTAAATATATACCAAAAGTGGCCCTGCTAGTATGTTTCTGTTGCCATCAACCCAT 3443
QY 3711 TACTTGTAGCTTAAACCAACATTAAGCTTATAGTCTGCGGATCAGAAATTCAGAAAT 3770
DB 3444 TACTTGTAGCTTAAACCAACATTAAGCTTATAGTCTGCGGATCAGAAATTCAGAAAT 3503
QY 3771 GGATGTCCTGAATGAATTAAGGTCAGAGAGCTGTCTCTTCTGAAGGCTTAG 3830
DB 3504 GGATGTCCTGAATGAATTAAGGTCAGAGAGCTGTCTCTTCTGAAGGCTTAG 3563
QY 3831 GGAGAAGCGGTTCTGTCATTTCAAGCTTCTAGAGCTGCTGATTCAGAGCTCA 3890
DB 3564 GGAGAAGCGGTTCTGTCATTTCAAGCTTCTAGAGCTGCTGATTCAGAGCTCA 3623
QY 3891 GTGGCTGTCAAGCTTTCTCATGATGCACTGTGACATGCGCCCTCCACTTCCCTC 3950
DB 3624 GTGGCTGTCAAGCTTTCTCATGATGCACTGTGACATGCGCCCTCCACTTCCCTC 3683
QY 3951 TTTGACTTAACAAAGCCCAAGAAATTCAGAAATTCCTCCATCTTAAATCTCTTA 4010
DB 3684 TTTGACTTAACAAAGCCCAAGAAATTCAGAAATTCCTCCATCTTAAATCTCTTA 3743
QY 4011 TCATCTGGAAGAGCTTTGTCATGCAAGCAACATAGCCACAGGTGGGATTAAGACC 4070
DB 3744 TCATCTGGAAGAGCTTTGTCATGCAAGCAACATAGCCACAGGTGGGATTAAGACC 3803
QY 4071 AGGACATCTTTGGGGTGTATTATCTGCTCAACAACCTTCTGCACTGACTCCACA 4130
DB 3804 AGGACATCTTTGGGGTGTATTATCTGCTCAACAACCTTCTGCACTGACTCCACA 3863
QY 4131 GGAAGGCTTAACAAATGATCTGGCGCACAGGATGTTTGTAGCTTGGCATCTTAAC 4190
DB 3864 GGAAGGCTTAACAAATGATCTGGCGCACAGGATGTTTGTAGCTTGGCATCTTAAC 3923
QY 4191 ACTTAAACCAACCCAGATCAGAAATCTGGCATGCTGGGCTCAATCTCACCTAGC 4250
DB 3924 ACTTAAACCAACCCAGATCAGAAATCTGGCATGCTGGGCTCAATCTCACCTAGC 3983
QY 4251 AACCACTGGCTGAGCTGGGCAACAAGCTTGTAGAAAGGAGTGTCACTTCAACAGG 4310
DB 3984 AACCACTGGCTGAGCTGGGCAACAAGCTTGTAGAAAGGAGTGTCACTTCAACAGG 4043
QY 4311 TCACCAAGCCCACTACGACCTTATCACTTCCCAATGAGGCTTAAGTGTGTTCTA 4370
DB 4044 TCACCAAGCCCACTACGACCTTATCACTTCCCAATGAGGCTTAAGTGTGTTCTA 4103
QY 4371 CTGATCAATGCCCCTGAGGTTGATTTATTTAATGAAGAAAGAACTGGGATTAATC 4430
DB 4104 CTGATCAATGCCCCTGAGGTTGATTTATTTAATGAAGAAAGAAAGTGGGATTAATC 4163
QY 4431 TCTAATCAGGTGATGACATGAGCAATGATGTGCTCACTTACCCTTTTCTTTT 4490
DB 4164 TCTAATCAGGTGATGATGAGCAATGATGTGCTCACTTACCCTTTTCTTTT 4223
QY 4491 TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 4550
DB 4224 TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 4283

QY 4551 TGAATGAGTGGGCAATCTCGGCTCACTGCAACCTTGCCTCTGGGCTCAAGCAAT 4610
DB 4284 TGAATGAGTGGGCAATCTCGGCTCACTGCAACCTTGCCTCTGGGCTCAAGCAAT 4343
QY 4611 CTCCCACTCAGCTTCCCAATATGTTGGATATCTGGCAACCAACATGCTCCAGCTTA 4670
DB 4344 CTCCCACTCAGCTTCCCAATATGTTGGATATCTGGCAACCAACATGCTCCAGCTTA 4403
QY 4671 TTTGTATTTTGTAGAGACAGGTTTCACATTTTCCAGGCTGGTCTCAACCTCT 4730
DB 4404 TTTGTATTTTGTAGAGACAGGTTTCACATTTTCCAGGCTGGTCTCAACCTCT 4463
QY 4731 GGGCTCAAGCAATCTCTGCTGCGCTCCCAAGTGTGGATTAAGATGAGCA 4790
DB 4464 GGGCTCAAGCAATCTCTGCTGCGCTCCCAAGTGTGGATTAAGATGAGCA 4523
QY 4791 CCGCATTCAGGCCCAACCTCTCATTTATACAAATTAAGTCTGCCAGTAATGAGCTTT 4850
DB 4524 CCGCATTCAGGCCCAACCTCTCATTTATACAAATTAAGTCTGCCAGTAATGAGCTTT 4583
QY 4851 GCTTCTCAACCCCTGCTGATCTGGAAGAGGATTAATGTTATAGCTTGTGAGCA 4910
DB 4584 GCTTCTCAACCCCTGCTGATCTGGAAGAGGATTAATGTTATAGCTTGTGAGCA 4643
QY 4911 GTCCCAATTAATTTCTGCGGCAAAACCTTCTTCAAAATTAATTAATTAATG 4970
DB 4644 GTCCCAATTAATTTCTGCGGCAAAACCTTCTTCAAAATTAATTAATTAATG 4703
QY 4971 TATTCATGAATTCATCTTGAATATGACCGCTCAACCTTGTTCATATGCAATTAATA 5030
DB 4704 TATTCATGAATTCATCTTGAATATGACCGCTCAACCTTGTTCATATGCAATTAATA 4763
QY 5031 AGGAATTTTAATAGTCTTAAATAGGCTGTACTGCAACCTTGAACATCTTCCAGAG 5090
DB 4764 AGGAATTTTAATAGTCTTAAATAGGCTGTACTGCAACCTTGAACATCTTCCAGAG 4823
QY 5091 GATAGATATTTAATAGTCAATGCTTGGCTTGGCTGATGAGCACTTCTCTGAAATG 5150
DB 4824 GATAGATATTTAATAGTCAATGCTTGGCTTGGCTGATGAGCACTTCTCTGAAATG 4883
QY 5151 TGGTCTGCGCAAGTACCTTGGCTTGTGAGCCGAGATGCTGACCTTGAATTAAGG 5210
DB 4884 TGGTCTGCGCAAGTACCTTGGCTTGTGAGCCGAGATGCTGACCTTGAATTAAGG 4943
QY 5211 CAAAGAGGCTGCGGCTTCTCTTCCCTCACTGAAGAGCCCTTATTTGAATTAATCACTGTG 5270
DB 4944 CAAAGAGGCTGCGGCTTCTCTTCCCTCACTGAAGAGCCCTTATTTGAATTAATCACTGTG 5003
QY 5271 GAGCCCTAGCCCTCAATCTGCAACTTCCCAACCTTCCAGGCCCTTCAAGCAGACTA 5330
DB 5004 GAGCCCTAGCCCTCAATCTGCAACTTCCCAACCTTCCAGGCCCTTCAAGCAGACTA 5063
QY 5331 GGTGCTGCAATTCACCCCAAGGTGGAATGGCTTCTTGAAGCTGACTTGTACCA 5390
DB 5064 GGTGCTGCAATTCACCCCAAGGTGGAATGGCTTCTTGAAGCTGACTTGTACCA 5123
QY 5391 TCACCGAATCACTGTTCCTGCAAGGACACCAAGTGGCAATTTCTTCAACTGAAGG 5450
DB 5124 TCACCGAATCACTGTTCCTGCAAGGACACCAAGTGGCAATTTCTTCAACTGAAGG 5183
QY 5451 TCAAAATCTCTGCAAGATGCTGCTCTGAGACAGATATTTCTGAGCTGTGCTCA 5510
DB 5184 TCAAAATCTCTGCAAGATGCTGCTCTGAGACAGATATTTCTGAGAGTGTGCTCA 5243
QY 5511 GTGAAGGAGCCAGCTGAGGAACCTGCTCTTTTCTTAAAGCCAGGCCCACTTAC 5570
DB 5244 GTGAAGGAGCCAGCTGAGGAACCTGCTCTTTTCTTAAAGCCAGGCCCACTTAC 5303
QY 5571 ATAAACATTTCAAGGTCACTGGAACAGTGAAGCCATTTGTGAAGCTTACGATG 5630
DB 5304 ATAAACATTTCAAGGTCACTGGAACAGTGAAGCCATTTGTGAAGCTTACGATG 5363
QY 5631 CCAGCCACTGCTATCCAGTGTCTGCATGCTTCAAGAGAAAGGCCAGGCCATGAGG 5690

```

Db      5364 CCGGCCACTGCTCATCCACGGTATGCGCTACAGGAGGCGCGCATGCAAG 5423
Qy      5691 ACTGCTCTTAATGCTGTGCTCATTCACAGAGGAGGAGGAGTCAAGTCAACT 5750
Db      5424 AATGGCTCTTAATGCTGTGCTCATTCACAGAGGAGGAGGAGTCAAGTCAACT 5483
Qy      5751 GGGACAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5810
Db      5484 GGGACAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5543
Qy      5811 GATCTCAGGCTTCACTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5870
Db      5544 GATCTCAGGCTTCACTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5603
Qy      5871 TATTCATTTCTGCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5930
Db      5604 TATTCATTTCTGCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5663
Qy      5931 AGTGTCTGCGCTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5990
Db      5664 AGTGTCTGCGCTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5723
Qy      5991 TGCAGGCTCAGATGCGCTCAGCAGATGCTGCTTAAATTTGATGATGATGATGATG 6050
Db      5724 TGCAGGCTCAGATGCGCTCAGCAGATGCTGCTTAAATTTGATGATGATGATG 5783
Qy      6051 AGGCTTTGGGATCAAGTCTTTCTCTTCAATTTGATGATGATGATGATGATGATGAT 6110
Db      5784 AGGCTTTGGGATCAAGTCTTTCTCTTCAATTTGATGATGATGATGATGATGATG 5843
Qy      6111 GTTTTGTCTCGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6163
Db      5844 GTTTTGTCTCGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5896

```

RESULT 3

AAH18243 ID AAH18243 standard; cDNA; 5573 BP.

XX AAH18243;

XX 26-JUN-2001 (first entry)

XX Human cDNA sequence SEQ ID NO:18190.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX BP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX MPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602.
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -

```

XX Claim 8; SEQ ID 18190; 2537P + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 5573 BP; 1373 A; 1524 C; 1290 G; 1386 T; 0 other;

```

Query Match 77.7%; Score 4786; DB 22; Length 5573;
Best Local Similarity 99.8%; Pred No. 0;
Matches 5566; Conservative 0; Mismatches 6; Indels 4; Gaps 4;

```

Qy 569 ATATTGACTGATATGTCAGGCGCCATGCTCTTTGGGAGGTTGAGCTTCAGCCCTGACAT 628
Db 1 ATATTGACTGATATGTCAGGCGCCATGCTCTTTGGGAGGTTGAGCTTCAGCCCTGACAT 60
Qy 629 CGTTGTTGCTTACCTCAACAGAACTTTGATGATGATGATGATGATGATGATGATG 688
Db 61 CGTTGTTGCTTACCTCAACAGAACTTTGATGATGATGATGATGATGATGATGATG 120
Qy 689 GTTTAGAGCTGAGTTTTCATCCCTCGCTGAGGAGATCGGTCCGGGTGAGAGTGCC 748
Db 121 G-TTAGAGCTGAGTTTTCATCCCTCGCTGAGGAGATCGGTCCGGGTGAGAGTGCC 179
Qy 749 CAGACGAGTCACTACTCCATCAGCGGCGGATGATGATGATGATGATGATGATGATG 808
Db 180 CAGACGAGTCACTACTCCATCAGCGGCGGATGATGATGATGATGATGATGATGATG 239
Qy 809 CCTTCTGACGATGAGCACTGTGTCCGGATCAAGTCAAGAGAGAGTGAATAGGCT 868
Db 240 CCTTCTGACGATGAGCACTGTGTCCGGATCAAGTCAAGAGAGAGTGAATAGGCT 299
Qy 869 TACACCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 928
Db 300 TACACCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 359
Qy 929 CTATTAAGAGTCTGATCATGATGATGATGATGATGATGATGATGATGATGATGATG 988
Db 360 CTATTAAGAGTCTGATCATGATGATGATGATGATGATGATGATGATGATGATGATG 419
Qy 989 TGTCTGCAATACCCAGAGGCTTCCCTGAGAGTGAAGTCAATGACGTGCAAGTTGTG 1048
Db 420 TGTCTGCAATACCCAGAGGCTTCCCTGAGAGTGAAGTCAATGACGTGCAAGTTGTG 479
Qy 1049 TTCTGCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1108
Db 480 TTCTGCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 539
Qy 1109 GGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1168
Db 540 GGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 599
Qy 1169 AGCTGAGAGCAAGAGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1228

```


Db 600 AGCTGAGAGCAAGAGCGCTGGGAGACATGCGGGGAACTTCAACCTCTCTGCAAGGCT 659
Qy 1229 GTGACCAAGATGCGCCAAAGTCCAGGGATCTCCGGCTGCAAGTTTCAAGTTTGGTCAAC 1288
Db 660 GTGACCAAGATGCGCCAAAGTCCAGGGATCTCCGGCTGCAAGTTTCAAGTTTGGTCAAC 719
Qy 1289 ATTCACAAAATGAAAGCAATAAATCTACGTGGTTGACTTGAATTAAGACGACATGT 1348
Db 720 ATTCACAAAATGAAAGCAATAAATCTACGTGGTTGACTTGAATTAAGACGACATGT 779
Qy 1349 CACTCAACATCCAGGCCAGGGCCCGTCAACAGAGCCGCAAGTTTGTCCCTGGCTGTTTGG 1408
Db 780 CACTCAACATCCAGGCCAGGGCCCGTCAACAGAGCCGCAAGTTTGTCCCTGGCTGTTTGG 839
Qy 1409 TGTGTTAGAAATCTCGGACCTGAGTAGAGAACTTCACTGACATCTGGCTCAAAACA 1468
Db 840 TGTGTTAGAAATCTCGGACCTGAGTAGAGAACTTCACTGACATCTGGCTCAAAACA 899
Qy 1469 AAATCTCTCTCTTGTGATGATCTGACAGCTCTGTGATGAAATGTGAAAAAACATTA 1528
Db 900 AAATCTCTCTCTTGTGATGATCTGACAGCTCTGTGATGAAATGTGAAAAAACATTA 959
Qy 1529 GCTGCACAGACCAACCGGTACTGCAAAAGAAATCTTACTGACTCCAGGTGCCAGTAGCA 1588
Db 960 GCTGCACAGACCAACCGGTACTGCAAAAGAAATCTTACTGACTCCAGGTGCCAGTAGCA 1019
Qy 1589 TCCTCCACCTGCTGTGAGCTGACATGATCTTCCCTCGAAGGCTGCTGGTCCCAAGACA 1648
Db 1020 TCCTCCACCTGCTGTGAGCTGACATGATCTTCCCTCGAAGGCTGCTGGTCCCAAGACA 1079
Qy 1649 GGTCAAGCTGTGCTGTGCTGACAGCCAGAAAGCTGACAGACATACACAGAAAGCCCT 1708
Db 1080 GGTCAAGCTGTGCTGTGCTGACAGCCAGAAAGCTGACAGACATACACAGAAAGCCCT 1139
Qy 1709 GCAACACCAAGCTTCAAGTACTGCTGTGCGCAAGTCCATCCAGCCAGACCTGTACTTG 1768
Db 1140 GCAACACCAAGCTTCAAGTACTGCTGTGCGCAAGTCCATCCAGCCAGACCTGTACTTG 1199
Qy 1769 GCTCCTTCGCGCGGAGGCTCTATCAGACAGATCCAGGTGAAGCAAGCAATCTCGGTGA 1828
Db 1200 GCTCCTTCGCGCGGAGGCTCTATCAGACAGATCCAGGTGAAGCAAGCAATCTCGGTGA 1259
Qy 1829 CCTTCGCACTTTTGGCCCGAGCTTCCAAAGAGGCTTCCAGGCGAGGTCTGACGAGTGT 1888
Db 1260 CCTTCGCACTTTTGGCCCGAGCTTCCAAAGAGGCTTCCAGGCGAGGTCTGACGAGTGT 1319
Qy 1889 CCTTTATACCTTATTTCAAAAGAGAGGCGTTTTCACGCTGACCCCTGACACAAAAAGCA 1948
Db 1320 CCTTTATACCTTATTTCAAAAGAGAGGCGTTTTCACGCTGACCCCTGACACAAAAAGCA 1379
Qy 1949 AGGTCTACCTGAGAGACCCCAACTGGGACCGGGGCTGCAATCCCTCACTCTGTGTCT 2008
Db 1380 AGGTCTACCTGAGAGACCCCAACTGGGACCGGGGCTGCAATCCCTCACTCTGTGTCT 1439
Qy 2009 GGAACATCAAGCTGCGCCAGAGACAGGTGAGCTGCTGACTTCTTTAAGAGCGAGCG 2068
Db 1440 GGAACATCAAGCTGCGCCAGAGACAGGTGAGCTGCTGACTTCTTTAAGAGCGAGCG 1499
Qy 2069 GCGTGTCTGCGACAGAGGCGCGCATTTATATCAATCCAGAGACGCGGAGCCCGGCTG 2128
Db 1500 GCGTGTCTGCGACAGAGGCGCGCATTTATATCAATCCAGAGAGCGGAGCCCGGCTG 1559
Qy 2129 AGGAATCTTCAGGCTGCGACAGAGATGTGTCTCCCAAGGCAAGCTTCCACATCAAGCT 2188
Db 1560 AGGAATCTTCAGGCTGCGACAGAGATGTGTCTCCCAAGGCAAGCTTCCACATCAAGCT 1619
Qy 2189 TCTGGGTCAACATCTCTAAGTCCAGCCCAAGAGCGGCAAGAGCTAGAGCTGCTCTTCT 2248
Db 1620 TCTGGGTCAACATCTCTAAGTCCAGCCCAAGAGCGGCAAGAGCTAGAGCTGCTCTTCT 1679
Qy 2249 CGGTGACCTTACCCCAAGAGCTGTGACTGTCACTCTCATCGGACGAGGTGGAG 2308

Db 1680 CGGTGACCTTACCCCAAGAGCTGTGACTGTGATCTCTCATCTCCAGCGGTGGAG 1739
Qy 2309 GTGAGTCTTACTGCTGTCTGCTCGGCTCATCATTTGTGTGAAAAAGAAATA 2368
Db 1740 GTGAGTCTTACTGCTGTCTGCTCGGCTCATCATTTGTGTGAAAAAGAAATA 1799
Qy 2369 AGAAGACAAACAAAGGCCCCGCTGTGGTATCTACAAATGGCAACATCAATCTGAGATGC 2428
Db 1800 AGAAGACAAACAAAGGCCCCGCTGTGGTATCTACAAATGGCAACATCAATCTGAGATGC 1859
Qy 2429 CGAGGACGCAAAAAGTTTTCAGAAAGGCGCAAGACATACATCTCCATGTGTATGACAG 2488
Db 1860 CGAGGACGCAAAAAGTTTTCAGAAAGGCGCAAGACATACATCTCCATGTGTATGACAG 1919
Qy 2489 TCATGAGGACACATATGATATATGGGCAATGTGCTACAGGATTCAGCGGCTCTTCCTGC 2548
Db 1920 TCATGAGGACACATATGATATATGGGCAATGTGCTACAGGATTCAGCGGCTCTTCCTGC 1979
Qy 2549 AGCCAGAGTGGACACATACCGGCGGTTCCAGGGACCAATGGGGGTCTGTCTCCCTCC 2608
Db 1980 AGCCAGAGTGGACACATACCGGCGGTTCCAGGGACCAATGGGGGTCTGTCTCCCTCC 2039
Qy 2609 CACCCACATATGCTTCAGGGCCCCCACTGCAAAAGTTGGCCACTGAGAGCCACTCTC 2668
Db 2040 CACCCACATATGCTTCAGGGCCCCCACTGCAAAAGTTGGCCACTGAGAGCCACTCTC 2099
Qy 2669 GCTCCCTCTGAGCTGAGATGAGTAAACGTTACCTTCCATCCCAACATGGGGATG 2728
Db 2100 GCTCCCTCTGAGCTGAGATGAGTAAACGTTACCTTCCATCCCAACATGGGGATG 2159
Qy 2729 TAAGCAGAAAGACACAGACATTCCTTACTGAAACACTCAGAGGCCATGAGCCAGCAG 2788
Db 2160 TAAGCAGAAAGACACAGACATTCCTTACTGAAACACTCAGAGGCCATGAGCCAGCAG 2219
Qy 2789 AATAACTGATCATTCATTCAGAGCGCTTGTGCTGATTTCAATAAGCAGGCACTGAGACAC 2848
Db 2220 AATAACTGATCATTCATTCAGAGCGCTTGTGCTGATTTCAATAAGCAGGCACTGAGACAC 2279
Qy 2849 GGTCCGTCTCTTACCAAGAAATCTTAAGAAAGAAATTTATAGAAAGAAACACACAGA 2908
Db 2280 GGTCCGTCTCTTACCAAGAAATCTTAAGAAAGAAATTTATAGAAAGAAACACACAGA 2339
Qy 2909 GGTTCCTCGACACCGGCAACTTCACTGAGTGTGCTGAGTCACTTCTAAGGCAAGACA 2968
Db 2340 GGTTCCTCGACACCGGCAACTTCACTGAGTGTGCTGAGTCACTTCTAAGGCAAGACA 2399
Qy 2969 TTGAATAATGATGAATTCGAATCTGAGATACAGTCAAGCTCATGTGCTCTCAACTTA 3028
Db 2400 TTGAATAATGATGAATTCGAATCTGAGATACAGTCAAGCTCATGTGCTCTCAACTTA 2459
Qy 3029 GGTGTGTGGTTAGCCACGCTGTATATGAGAGAGAGGCTGAGTCACTTAGCACTAAGG 3088
Db 2460 GGTGTGTGGTTAGCCACGCTGTATATGAGAGAGAGGCTGAGTCACTTAGCACTAAGG 2519
Qy 3089 TTGCAGCAAGCCCTGAGTTTCAAGAGTGTAAACAGAGGCTTGCCTCTTCAAGACAAAGT 3148
Db 2520 TTGCAGCAAGCCCTGAGTTTCAAGAGTGTAAACAGAGGCTTGCCTCTTCAAGACAAAGT 2579
Qy 3149 TCCAAATTCGAAGAGCTTACCTGAGGCTCTACTCTCACTGAGGCTCCACAGATGAATAAC 3208
Db 2580 TCCAAATTCGAAGAGCTTACCTGAGGCTCTACTCTCACTGAGGCTCCACAGATGAATAAC 2639
Qy 3209 GACAAATGTGCTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 3268
Db 2640 GACAAATGTGCTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2699
Qy 3269 GTATTAACCACTAGCTCTTTTCACTGACTTATTAATACTAATCTAATCTAGTGTGAT 3328
Db 2700 GTATTAACCACTAGCTCTTTTCACTGACTTATTAATACTAATCTAATCTAGTGTGAT 2759
Qy 3329 GCTGTGGTTGTGACTTCTTACTGACCGCTAGATTAACGCTGTGCTCCAGAGTGTGG 3388
Db 2760 GCTGTGGTTGTGACTTCTTACTGACCGCTAGATTAACGCTGTGCTCCAGAGTGTGG 2819

QY 3389 GAATTAATTACAACTGTCCAAACAGAAAAGAAATGTGTGTTTGAGCAGCATTTGACACA 3448
DB 2820 GAATTAATTACAACTGTCCAAACAGAAAAGAAATGTGTGTTTGAGCAGCATTTGACACA 2879
QY 3449 TAATCTGTTTGAATGAAGACTTCTGATTTCTTAGTCGGTGTGTGTTATCCCATTTG 3508
DB 2880 TAATCTGTTTGAATGAAGACTTCTGATTTCTTAGTCGGTGTGTGTTATCCCATTTG 2939
QY 3509 GAATTCATCTTGAATTCAGATTGTCTATAGTCTTAGCAATAAGAAAATTTTCTCAAGT 3568
DB 2940 GAATTCATCTTGAATTCAGATTGTCTATAGTCTTAGCAATAAGAAAATTTTCTCAAGT 2999
QY 3569 TTCCATGTGCGGTTCTCTAGCTGACAGCAATCTTGAATTTTAAAGAAAATTTTGAAG 3628
DB 3000 TTCCATGTGCGGTTCTCTAGCTGACAGCAATCTTGAATTTTAAAGAAAATTTTGAAG 3059
QY 3629 ATATTTCTCATCTCTTAAATATTTTAAATATATATATATATATATATATATATATATAT 3688
DB 3060 ATATTTCTCATCTCTTAAATATTTTAAATATATATATATATATATATATATATATATAT 3119
QY 3689 TTCTGTGTCACATGCAACCCATTACTGTAGCTTAAACACATTTAGTCTTATAGTC 3748
DB 3120 TTCTGTGTCACATGCAACCCATTACTGTAGCTTAAACACATTTAGTCTTATAGTC 3179
QY 3749 CTGCGGATCAGAAATTCAAAATGATGTCCCTGAATGAATAATCAGGTGTACAGAGCT 3808
DB 3180 CTGCGGATCAGAAATTCAAAATGATGTCCCTGAATGAATAATCAGGTGTACAGAGCT 3239
QY 3809 GTGCTCTCTTGAAGGCTCTAGGAGAGAGCGGTTCTTGCCATTTCAAGCTTCTAGAG 3868
DB 3240 GTGCTCTCTTGAAGGCTCTAGGAGAGAGCGGTTCTTGCCATTTCAAGCTTCTAGAG 3299
QY 3869 CTGCGTCATTTCCAGGCTCCAGTGTGCTGATCAAGCTTTTCTCATATGATCATCTGTA 3928
DB 3300 CTGCGTCATTTCCAGGCTCCAGTGTGCTGATCAAGCTTTTCTCATATGATCATCTGTA 3359
QY 3929 CACTGCGCTTCCCACTTCCCTCTTTTGAATTTAAAGGCCACAGAAAGATCCAGATTAAT 3988
DB 3360 CACTGCGCTTCCCACTTCCCTCTTTTGAATTTAAAGGCCACAGAAAGATCCAGATTAAT 3419
QY 3989 CTCTCCATCTAAAGATTCCTTCAATCTCTGAAAGAGCTTTTGCCATGACAGACATTA 4048
DB 3420 CTCTCCATCTAAAGATTCCTTCAATCTCTGAAAGAGCTTTTGCCATGACAGACATTA 3479
QY 4049 GCCACAGGTGGGATTTAGACACAGACATCTTTGGGGTGTCTTTATTTCTGCTTACCAC 4108
DB 3480 GCCACAGGTGGGATTTAGACACAGACATCTTTGGGGTGTCTTTATTTCTGCTTACCAC 3539
QY 4109 CTTCCTGCACTGACTCCACAGAGAGGCTACAAAATGATCTGGCGCACAGGGATGTTT 4168
DB 3540 CTTCCTGCACTGACTCCACAGAGAGGCTACAAAATGATCTGGCGCACAGGGATGTTT 3599
QY 4169 TGTTTAGCTTGGCGACTTAAACATT-AAAAAAACCCCAATCAGAAAGATCTGGCCATG 4227
DB 3600 TGTTTAGCTTGGCGACTTAAACATTAAAAAAACCCCAATCAGAAAGATCTGGCCATG 3659
QY 4228 TGGGGCTCATCTTCTCAGCAACATCTGCTGAGCTGGGCAACAGCTTGCCTTTA 4287
DB 3660 TGGGGCTCATCTTCTCAGCAACATCTGCTGAGCTGGGCAACAGCTTGCCTTTA 3719
QY 4288 GAAAGGGGTGTCACTTCCACAGGTCAACAGGCCACACTACGCCCTATCATCTTCCACA 4347
DB 3720 GAAAGGGGTGTCACTTCCACAGGTCAACAGGCCACACTACGCCCTATCATCTTCCACA 3779
QY 4348 ATGAGGTAGTGTGTTTCTTACTGATCAATGCCCCCTGCAAGTTGATTTATTTGAATG 4407
DB 3780 ATGAGGTAGTGTGTTTCTTACTGATCAATGCCCCCTGCAAGTTGATTTATTTGAATG 3839
QY 4408 AAAAAGAAAGTGGGATTAATCTTAATCAGGTAGTAGACCATAGAACATGTGTG 4467
DB 3840 AAAAAGAAAGTGGGATTAATCTTAATCAGGTAGTAGACCATAGAACATGTGTG 3899

QY 4468 TCACATTAACCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 4527
DB 3900 TCACATTAACCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 3959
QY 4528 GGATCTCATTTCTGTTGCTCAGGCTGAGAGTGAAGGCGCAATCTGGGCTCATCTGCACT 4587
DB 3960 GGATCTCATTTCTGTTGCTCAGGCTGAGAGTGAAGGCGCAATCTGGGCTCATCTGCACT 4019
QY 4588 CTGCTCTGGGCTCAGCAATTTCTCCACCTGACCTTCCCAATATGCTGGATCATCTGG 4647
DB 4020 CTGCTCTGGGCTCAGCAATTTCTCCACCTGACCTTCCCAATATGCTGGATCATCTGG 4079
QY 4648 CACAAACCACTGCGCCAGCTAATTTTGTATTTTGTATGAGACAGGTTTACCATGTT 4707
DB 4080 CACAAACCACTGCGCCAGCTAATTTTGTATTTTGTATGAGACAGGTTTACCATGTT 4138
QY 4708 GCCAGGCTGCTCAACCTCTGCGGCTCAGCAATTCCTCTGCTCCGCTCCGAGCT 4767
DB 4139 GCCAGGCTGCTCAACCTCTGCGGCTCAGCAATTCCTCTGCTCCGCTCCGAGCT 4198
QY 4768 GCTGGGATTTACAGATGTGAGCCACCGCATCCAGCCCAACCCCTCATTTATACCAATTAC 4827
DB 4199 GCTGGGATTTACAGATGTGAGCCACCGCATCCAGCCCAACCCCTCATTTATACCAATTAC 4258
QY 4828 CTGCCAGTAACTGTGACATTTTCTTCTCTCAACCCCTGCTGTGATCTGAGAGAGGGA 4887
DB 4259 CTGCCAGTAACTGTGACATTTTCTTCTCTCAACCCCTGCTGTGATCTGAGAGAGGGA 4318
QY 4888 TTATGTTATAGCTGTGACAGCAGTCCCAAGTTCAATATTTCTGGGGCAAAAATCTCTT 4947
DB 4319 TTATGTTATAGCTGTGACAGCAGTCCCAAGTTCAATATTTCTGGGGCAAAAATCTCTT 4378
QY 4948 CAAAAATTAATGTACTTATTTATTTCAATGAATTTCACTTGAATATGACCGCTCAA 5007
DB 4379 CAAAAATTAATGTACTTATTTATTTCAATGAATTTCACTTGAATATGACCGCTCAA 4438
QY 5008 CTGTGTCATGTCATTAATGAAGAAATTTTATGCTCTCTAAATGCGGTACTGCA 5067
DB 4439 CTGTGTCATGTCATTAATGAAGAAATTTTATGCTCTCTAAATGCGGTACTGCA 4498
QY 5068 GACCTCTTGAACATTTCCAGAGATAGGATTTTATAGCATAGCCCTGGGCTTGCTAT 5127
DB 4499 GACCTCTTGAACATTTCCAGAGATAGGATTTTATAGCATAGCCCTGGGCTTGCTAT 4558
QY 5128 GGCACCTTTCCTTCTGAAAGCTGTGCTCTGCCCCAGTGAACCTTGTGAGCGCA 5187
DB 4559 GGCACCTTTCCTTCTGAAAGCTGTGCTCTGCCCCAGTGAACCTTGTGAGCGCA 4618
QY 5188 GATGCTGACCTGCAATAAGGCGCAAGAGGCGCTGCGCTTCTTCCCTCATCTGAGAG 5247
DB 4619 GATGCTGACCTGCAATAAGGCGCAAGAGGCGCTGCGCTTCTTCCCTCATCTGAGAG 4678
QY 5248 CCTTATTTGAATTAACGTGTGAGGCGCTTACCTCATCTTGTGCAATTTCCCAACCTC 5307
DB 4679 CCTTATTTGAATTAACGTGTGAGGCGCTTACCTCATCTTGTGCAATTTCCCAACCTC 4738
QY 5308 CCAGCCCTTCCAAACAGGACTAGGTGCTGCAATTTCAACCCAAAGTGGATTTGCTTC 5367
DB 4739 CCAGCCCTTCCAAACAGGACTAGGTGCTGCAATTTCAACCCAAAGTGGATTTGCTTC 4797
QY 5368 CTTAAGCTGCTACTTGTGACCATCAACGACATCACTGTGCTGCAAGACACACGTC 5427
DB 4798 CTTAAGCTGCTACTTGTGACCATCAACGACATCACTGTGCTGCAAGACACACGTC 4857
QY 5428 GGCATTTTCTTCACTAGAGGCTCAAAATCTCTGAGCAAGTTGCTGCTGAGACCA 5487
DB 4858 GGCATTTTCTTCACTAGAGGCTCAAAATCTCTGAGCAAGTTGCTGCTGAGACCA 4917
QY 5488 GTATTTCTGAGAGCTGCTCATGTAAGGGGCCAGCCTGAGAAACCTGTGCTTTTC 5547
DB 4918 GTATTTCTGAGAGCTGCTCATGTAAGGGGCCAGCCTGAGAAACCTGTGCTTTTC 4977
QY 5548 TTAAAGCCAGGCCCACTTACATTAACATTTTCAAGGTCACTGAAAACACTGAAGTGC 5607

DB 4978 TTTAAAGCCAGGCCCCCACTTACATTAACATTTTCAGGGTCACTGAAACAGTGAAGTGC 5037
OY 5608 CATTGTTGAGGCTTCTGATGATCCAGCCCACTGCTCATCCAGTGTGCGCATGCCCA 5667
DB 5038 CATTGTTGAGGCTTCTGATGATCCAGCCCACTGCTCATCCAGTGTGCGCATGCCCA 5097
OY 5668 CGAGGAAGGCGAGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5727
DB 5098 CGAGGAAGGCGAGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5157
OY 5728 AAGGTTCTCAAGGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 5787
DB 5158 AAGGTTCTCAAGGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 5217
OY 5788 AAGGTTCTCAAGGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 5847
DB 5218 AAGGTTCTCAAGGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 5277
OY 5848 CAGATACCTTTTAAAGTGGATTTTAACTTTCTGCTGCTTTAAATTAATTAATTAATTA 5907
DB 5278 CAGATACCTTTTAAAGTGGATTTTAACTTTCTGCTGCTTTAAATTAATTAATTAATTA 5337
OY 5908 AAAAAACAAGTATCAAGCTTTTAAAGTCTTCTGCTGCTTTAAATTAATTAATTAATTA 5967
DB 5338 AAAAAACAAGTATCAAGCTTTTAAAGTCTTCTGCTGCTTTAAATTAATTAATTAATTA 5397
OY 5968 CCGGGTTCCTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 6027
DB 5398 CCGGGTTCCTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 5457
OY 6028 ATTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6087
DB 5458 ATTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5517
OY 6088 GCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6143
DB 5518 GCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5573

RESULT 4
AB211770
ID AB211770 standard; cDNA; 2669 BP.
XX
AC AB211770;
XX
DT 20-JAN-2003 (first entry)
XX
DE Human polynucleotide SEQ ID NO 652.
XX
XX Human; genome mapping; gene therapy; food supplement; virus; fungus;
cell-proliferative disorder; neurodegenerative disease; bacterial;
Parkinson's disease; Alzheimer's disease; autoimmune disease;
multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
arthritis; cystic fibrosis; immunomodulator; neoplastic; neuroprotective;
anti-Parkinsonian; antidiabetic; immunosuppressive; dermatological;
haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
antiarthritic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200270539-A2.
XX
PD 12-SEP-2002.
XX
XX 05-MAR-2002; 2002WO-US05095.
XX
PR 05-MAR-2001; 2001US-0799451.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F,
Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;

PI Wehman T, Wang J, Wang D, Drmanac RT;
XX WPI; 2002-759812/82.
DR P-PSDB; ABP69553.
XX
PT New polynucleotides comprising sequences assembled from expressed
sequence tags (ESTs), useful for treating cell-proliferative,
neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or
platelet or coagulation disorders -
XX
PS Claim 1; SEQ ID NO 652; 1012pp + Sequence Listing; English.
XX
XX The invention relates to an isolated polynucleotide (1) comprising a
nucleotide sequence selected from any of 948 sequences
(AB211119-AB212066) or their mature protein coding portion, active domain
coding protein or complementary sequences. The polynucleotides are useful
for identifying expressed genes or for physical mapping of human genome.
The encoded polypeptides (ABP68902-ABP6949) are useful as molecular
weight markers, as a food supplement, for generating antibodies, in
medical imaging, screening and diagnostic assays and for treating
cell-proliferative disorders (cancer), neurodegenerative diseases
(Parkinson's or Alzheimer's disease), autoimmune diseases (multiple
sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid
disorders, platelet or coagulation disorders, wound, burns, incision,
ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,
CC parasitic), arthritis, etc.
CC Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2669 BP; 634 A; 790 C; 675 G; 570 T; 0 other;
Query Match 41.0%; Score 2524; DB 24; Length 2669;
Best Local Similarity 99.9%; Pred. No. 0; Mismatches 2; Indels 0; Gaps 0;
Matches 2624; Conservative 0; P-PSDB; ABP69553.
OY 219 GGGGCGCTGGGGCGCGCGCTCCCACTGCTTTTCCCACTGAGGCGGCGGCGG 278
DB 44 GGGGCGCTGGGGCGCGCGCTCCCACTGCTTTTCCCACTGAGGCGGCGGCGG 103
OY 279 AGTCATGGCGCGCGCTGAACTGCGGGGCTCTATGCACTGCTAGGGCTTCTGCTGG 338
DB 104 AGTCATGGCGCGCGCTGAACTGCGGGGCTCTATGCACTGCTAGGGGCTTCTGCTGG 163
OY 339 TGGCGGCG 398
DB 164 TGGCGGCG 223
OY 399 CATTACAGTTCATTAAGCTGGGAGCCCGACTCTGCTGGCAAAACCTGTTACATGCT 458
DB 224 CATTACAGTTCATTAAGCTGGGAGCCCGACTCTGCTGGCAAAACCTGTTACATGCT 283
OY 459 CATTCTTAAGAGCATATTAACCATGTTGTCATCAAGTCTGGAGAAAGATAGTCTTAC 518
DB 284 CATTCTTAAGAGCATATTAACCATGTTGTCATCAAGTCTGGAGAAAGATAGTCTTAC 343
OY 519 CTTTACGTCGAGAGTCTGAGAAATCACTTTGTCTATAGATCCGAAATTAATTAAGTCTG 578
DB 344 CTTTACGTCGAGAGTCTGAGAAATCACTTTGTCTATAGATCCGAAATTAATTAAGTCTG 403
OY 579 TATGTCAGGCGCATGCTTTTGGGAGGTTTCACTGAGGCTTCACTGAGGCTTCACTGAGG 638
DB 404 TATGTCAGGCGCATGCTTTTGGGAGGTTTCACTGAGGCTTCACTGAGGCTTCACTGAGG 463
OY 639 TACCTCAACAGAACTTTTCACTGAGGATGTCAGAACTATTAAGAGCATGCTGTTAGAGCT 698
DB 464 TACCTCAACAGAACTTTTCACTGAGGATGTCAGAACTATTAAGAGCATGCTGTTAGAGCT 523
OY 699 GCAAGTTTCAATCCCTGCGCTGAGGAGATGCTGCGGCTGAGGAGGCTGAGGAGGAGT 758
DB 524 GCAAGTTTCAATCCCTGCGCTGAGGAGATGCTGCGGCTGAGGAGGCTGAGGAGGAGT 583
OY 759 CACTCACTCCATCAAGCGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 818

Db 584 CACTCCTCATAGGCGCGAATCGATGCGACCGGTGTCAGGATGGAACCTTCTGACG 643
Qy 819 CAATGCACTGTGTCCCGATCAAGATGCAAGAAGGTGAAATGAGCTTACACCTCC 878
Db 644 CAATGCACTGTGTCCCGATCAAGATGCAAGAAGGTGAAATGAGCTTACACCTCC 703
Qy 879 ATGGTTCCACCCAGAAATGTCTCCGCTTACGATTTGCAACCCGCTCATCTATAAAG 938
Db 704 ATGGTTCCACCCAGAAATGTCTCCGCTTACGATTTGCAACCCGCTCATCTATAAAG 763
Qy 939 TCTGTGATCATGAGTCTGTGTTTGAAGGTGAAAGGTGAGCAACCTGATGTCTGCAA 998
Db 764 TCTGTGATCATGAGTCTGTGTTTGAAGGTGAAAGGTGAGCAACCTGATGTCTGCAA 823
Qy 999 CTACCCAGAAAGCTTCCCTGAGATGAGTCATGACGTGACAGTTTGTCTGCTGACA 1058
Db 824 CTACCCAGAAAGCTTCCCTGAGATGAGTCATGACGTGACAGTTTGTCTGCTGACA 883
Qy 1059 CTTGCGGGCCAGCGTCTCTTCTCTCAACCTTCAACTGTGAGAGAGAGAG 1118
Db 884 CTTGCGGGCCAGCGTCTCTTCTCTCAACCTTCAACTGTGAGAGAGAGAG 943
Qy 1119 GCGGGTTGAATATTAATCCCGGGCTCCAGCAACCCGAGGTGTTCAAGCTGAGAG 1178
Db 944 GCGGGTTGAATATTAATCCCGGGCTCCAGCAACCCGAGGTGTTCAAGCTGAGAG 1003
Qy 1179 CAAGCAAGCTGGGAACATGCGGGGAACTTCAACTCTCTCTGCAAGGCTGTGACAGA 1238
Db 1004 CAAGCAAGCTGGGAACATGCGGGGAACTTCAACTCTCTCTGCAAGGCTGTGACAGA 1063
Qy 1239 TGCCCAAGTCCAGGATCTCCGGCTGCAAGTTCCAAAGTTTGTGTCACATCCACAAA 1298
Db 1064 TGCCCAAGTCCAGGATCTCCGGCTGCAAGTTTGTGTCACATCCACAAA 1123
Qy 1299 TGAAGAAATAAATCTAGTGTGTTGACTTGAATAGCGAGCCATGCACTCAACAT 1358
Db 1124 TGAAGAAATAAATCTAGTGTGTTGACTTGAATAGCGAGCCATGCACTCAACAT 1183
Qy 1359 CGAGCAAGCGCCCGTCAAAAGAGCGCAAGTTTGTCTGCTGCTGTGTCTAGA 1418
Db 1184 CGAGCAAGCGCCCGTCAAAAGAGCGCAAGTTTGTCTGCTGCTGTGTCTAGA 1243
Qy 1419 ATCTCGGACTGCAATGCAACCTCACTGCAATCTGCTCAACCAAAATCTCTT 1478
Db 1244 ATCTCGGACTGCAATGCAACCTCACTGCAATCTGCTCAACCAAAATCTCTT 1303
Qy 1479 CTTTTGATGATCTGACAGTCTGTGAGTGAATGTGGAATAAATCAATAGCTGACAGA 1538
Db 1304 CTTTTGATGATCTGACAGTCTGTGAGTGAATGTGGAATAAATCAATAGCTGACAGA 1363
Qy 1539 CCAACGGTACTGCAAAAGAAATCTTACTCTCAGAGTCCAGTGAACATCTCCACCT 1598
Db 1364 CCAACGGTACTGCAAAAGAAATCTTACTCTCAGAGTCCAGTGAACATCTCCACCT 1423
Qy 1599 GCGTGTGAGCTGCAATCTTCTCTGGAAGCTGTGTGTCGCAAGAGACAGGCTCACT 1658
Db 1424 GCGTGTGAGCTGCAATCTTCTCTGGAAGCTGTGTGTCGCAAGAGACAGGCTCACT 1483
Qy 1659 GGTGCTGTGAGCCCAAGAAAGTGTGAGAGCATACAGAGAAAGCCCTGCAACACAG 1718
Db 1484 GGTGCTGTGAGCCCAAGAAAGTGTGAGAGCATACAGAGAAAGCCCTGCAACACAG 1543
Qy 1719 CTTACAGTACTCTGTGAGCAGTGCATACCCAGCAGAGACCTGTACTCGGCTCTTCTG 1778
Db 1544 CTTACAGTACTCTGTGAGCAGTGCATACCCAGCAGAGACCTGTACTCGGCTCTTCTG 1603
Qy 1779 CCGGGAAGGCTCTATCAAGCAGATCCAGGTGAAGCAAAATCTCGGTGACCTTTCGAC 1838
Db 1604 CCGGGAAGGCTCTATCAAGCAGATCCAGGTGAAGCAAAATCTCGGTGACCTTTCGAC 1663
Qy 1839 CTTTGCCCCAGCTTCCAAAGAGGCTTCAGGCAAGGCTTGAAGGCTCTTATATACC 1898

Db 1664 CTTTGCCCCAGCTTCCAAAGAGGCTTCAGGCAAGGCTTGAAGGCTCTTATATACC 1723
Qy 1899 TTATTTCAAGAGGAAGCGTTTTCACGGTGAACCCCTGACACAAAGCAAGTCTACT 1958
Db 1724 TTATTTCAAGAGGAAGCGTTTTCACGGTGAACCCCTGACACAAAGCAAGTCTACT 1783
Qy 1959 GAGGACCCCACTGGGACCCGGGCTGCTCATCTTCTGCTGTGTCTTGAATAATCAG 2018
Db 1784 GAGGACCCCACTGGGACCCGGGCTGCTCATCTTCTGCTGTGTCTTGAATAATCAG 1843
Qy 2019 CGTGCCCAAGAGCAAGGTGAGCTGCTGACTTCTTTTAAAGAGCGGAGCGGTGCTG 2078
Db 1844 CGTGCCCAAGAGCAAGGTGAGCTGCTGACTTCTTTTAAAGAGCGGAGCGGTGCTG 1903
Qy 2079 CCAGAAGGCGCGCATTCATGATCATCCAGAGCAAGCGAACCCGGGCTGAGAGATCTT 2138
Db 1904 CCAGAAGGCGCGCATTCATGATCATCCAGAGCAAGCGAACCCGGGCTGAGAGATCTT 1963
Qy 2139 CAGCTGAGACGAGATGTGCTTCCCAAGCCAGCTTCCACATCAAGCTTCTGGGTCAA 2198
Db 1964 CAGCTGAGACGAGATGTGCTTCCCAAGCCAGCTTCCACATCAAGCTTCTGGGTCAA 2023
Qy 2199 CATCTCTAATGACAGCCCAAGAGGCGCAAGAGCTGACCTGCTTCTCGGTGACAT 2258
Db 2024 CATCTCTAATGACAGCCCAAGAGGCGCAAGAGCTGACCTGCTTCTCGGTGACAT 2083
Qy 2259 TACCCCAAGACTGTGACTTGTACTCTCATCTGACAGCGGTGAGAGTGTCTT 2318
Db 2084 TACCCCAAGACTGTGACTTGTACTCTCATCTGACAGCGGTGAGAGTGTCTT 2143
Qy 2319 ACTGCTGTGCTGCTTCCGGGCTCATCTTGTGTGTGAAAAAGAAAAAGAACAA 2378
Db 2144 ACTGCTGTGCTGCTTCCGGGCTCATCTTGTGTGTGAAAAAGAAAAAGAACAA 2203
Qy 2379 CAAGGGCCCCGCTGTGGGATCTAATAGGCAACATCAATAGTGAATGCGAGGAGCC 2438
Db 2204 CAAGGGCCCCGCTGTGGGATCTAATAGGCAACATCAATAGTGAATGCGAGGAGCC 2263
Qy 2439 AAAAAAGTTTCAGAAAGGCGAAGAGCAATGACTCCATGTGTATGCAAGTATCGAGA 2498
Db 2264 AAAAAAGTTTCAGAAAGGCGAAGAGCAATGACTCCATGTGTATGCAAGTATCGAGA 2323
Qy 2499 CACCATGTGATATGGGCACTGTGCTACAGATTCAGCGGCTCTTCTGCAAGCAAGGT 2558
Db 2324 CACCATGTGATATGGGCACTGTGCTACAGATTCAGCGGCTCTTCTGCAAGCAAGGT 2383
Qy 2559 GGAACCTTACCGGCGGTTCCAGGGCAACCATGGGGGTGTGTCTCCCTCCCAACCAT 2618
Db 2384 GGAACCTTACCGGCGGTTCCAGGGCAACCATGGGGGTGTGTCTCCCTCCCAACCAT 2443
Qy 2619 ATGCTTCAAGGGCCCAACTGCAAAAGTTGGCCACTGAGAGGACCACTCTGCTCCCTCC 2678
Db 2444 ATGCTTCAAGGGCCCAACTGCAAAAGTTGGCCACTGAGAGGACCACTCTGCTCCCTCC 2503
Qy 2679 TGAAGTCTGAGAGTGAACCGTACACTTCTCCATCTCCAAATGGGATGTAAAGCAGAA 2738
Db 2504 TGAAGTCTGAGAGTGAACCGTACACTTCTCCATCTCCAAATGGGATGTAAAGCAGAA 2563
Qy 2739 GGAACAAGACATTCCTTACTGAAACACTCAGGAGCCCATGAGCGAGAGAACTTGA 2798
Db 2564 GGAACAAGACATTCCTTACTGAAACACTCAGGAGCCCATGAGCGAGAGAACTTGA 2623
Qy 2799 TCCATTTCCAGAGCTTTGCTGAGTTTCAATAAGCAGGCACTGAGA 2844
Db 2624 TCCATTTCCAGAGCTTTGCTGAGTTTCAATAAGCAGGCACTGAGA 2669

RESULT 5
AAA26351
ID AAA26351 standard, cDNA, 2209 BP.
XX
AC
XX AAA26351:

29-JUN-2000 (first entry)
Human secreted protein gene 6 SEQ ID NO:16.
Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
antiHIV; antiinflammatory; nootropic; neuroprotective; antiallergic;
osteoporotic; antidiabetic; antibacterial; antidiabetic; antiashma;
antiproliferative; cardiac; gene therapy; cancer; neurological disorder;
immune disease; inflammation; blood disorder; tumour; ss.
Homo sapiens.
WO200006698-A1.
10-FEB-2000.
29-JUL-1999; 99WO-US17130.
30-JUL-1998; 98US-0094657.
05-AUG-1998; 98US-0095486.
06-AUG-1998; 98US-0095454.
06-AUG-1998; 98US-0095455.
12-AUG-1998; 98US-0096319.
(HUMA-) HUMAN GENOME SCI INC.
Komatsu GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y,
Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA,
Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenak M,
WPI; 2000-195282/17.
P-PsDB; AAY91456.
New isolated human genes and the secreted polypeptides they encode,
useful for diagnosis and treatment of e.g. cancers, neurological
disorders, immune diseases, inflammation or blood disorders -
Claim 1; Page 375; 634pb; English.
The polynucleotide sequences given in AAY91451 to AAY91459 encode the
human secreted proteins given in AAY91451 to AAY91459. The human secreted
proteins can have activities based on the tissues and cells they are
expressed in. Examples of the activities are: cytostatic;
immunosuppressive; antiHIV; antiinflammatory; nootropic; neuroprotective;
antiallergic; osteoporotic; antidiabetic; antibacterial; antidiabetic;
antiashma; antiproliferative; and cardiac. The polynucleotides and their
corresponding secreted proteins are useful for preventing, treating or
ameliorating medical conditions, e.g. by protein or gene therapy. Also
pathological conditions can be diagnosed by determining the amount of the
proteins in a sample or by determining the presence of mutations in the
polynucleotides. Specific uses are described for each of the
polynucleotides, based on which tissues they are most highly expressed
in, and include developing products for the diagnosis or treatment of
cancer, tumours, neurodegenerative disorders, developmental abnormalities
and foetal deficiencies, blood disorders, diseases of the immune system,
autoimmune diseases, hepatic and renal diseases, inflammation,
allergies, Alzheimer's and behavioural disorders, schizophrenia,
osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
cardiovascular disorders, reproductive disorders, gastrointestinal
disorders, respiratory disorders and metabolic disorders. The proteins
or polynucleotides can also be used as food additives or preservatives.
The proteins are also useful for identifying their binding partners.
AAY91451 to AAY91459 and AAY91450 are sequences used in the
exemplification of the present invention.
Sequence 2209 BP; 508 A; 646 C; 566 G; 477 T; 12 other;
Query Match 25.1%; Score 1545; DB 21; Length 2209;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 2145; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
204 AGCTGCGCGGCTTGGGCGCTGGGCGCGCTCCACCGTCTTTTCCACCGAG 263

2 AGCTGCGCGGCTTGGGCGCTGGGCGCGCTCCACCGTCTTTTCCACCGAG 61
GCCGAGGCGTCCCGAGCATGCGCGCTGAATCGCGGGCTCTATGCACTGCTAG 323
GCCGAGGCGTCCCGAGCATGCGCGCTGAATCGCGGGCTCTATGCACTGCTAG 121
GATTCTGCTGCTGGTGGCGCGCTCCCGCGCGGCGAGAGCTTTGAGATTGCTCT 383
GGTTCTGCTGCTGGTGGCGCGCTCCCGCGCGGCGAGAGCTTTGAGATTGCTCT 181
GCCACGAGAGAGCAATTAACAGTTCTATAAAGCTGGGAGACCCGACTTGTGCAAA 443
GCCACGAGAGAGCAATTAACAGTTCTATAAAGCTGGGAGACCCGACTTGTGCAAA 241
ACCGTGTACATGCTGCTATTAAGACATATTAACATGTTGCTCAATGCTGAGAGA 503
ACCGTGTACATGCTGCTATTAAGACATATTAACATGTTGCTCAATGCTGAGAGA 301
AAGATATGCTTTTACCTTTAGCTGCGAGATCTGAGATCACTTTGCTATGAGATCA 563
AAGATATGCTTTTACCTTTAGCTGCGAGATCTGAGATCACTTTGCTATGAGATCA 361
GAAATATGCTTTTACCTTTAGCTGCGAGATCTGAGATCACTTTGCTATGAGATCA 623
GAAATATGCTTTTACCTTTAGCTGCGAGATCTGAGATCACTTTGCTATGAGATCA 421
GAAATATGCTTTTACCTTTAGCTGCGAGATCTGAGATCACTTTGCTATGAGATCA 683
GAAATATGCTTTTACCTTTAGCTGCGAGATCTGAGATCACTTTGCTATGAGATCA 481
CATCGTTTAAAGCTGAGTTCATCCCTGCGCTGAGGAGATCGGTCCGGGAGAG 743
CATCGTTTAAAGCTGAGTTCATCCCTGCGCTGAGGAGATCGGTCCGGGAGAG 541
CATCGTTTAAAGCTGAGTTCATCCCTGCGCTGAGGAGATCGGTCCGGGAGAG 803
CTGCCAGAGAGAGTCACTCACTCCATGAGCGCGGAGATCGATGCTGAGAT 601
CTGCCAGAGAGAGTCACTCACTCCATGAGCGCGGAGATCGATGCTGAGAT 601
CGGAGAGAGAGTCACTCACTCCATGAGCGCGGAGATCGATGCTGAGAT 863
TGGAGAGAGAGTCACTCACTCCATGAGCGCGGAGATCGATGCTGAGAT 661
GAGCTTAAAGCTGAGTTCATCCCTGCGCTGAGGAGATCGGTCCGGGAGAG 923
GAGCTTAAAGCTGAGTTCATCCCTGCGCTGAGGAGATCGGTCCGGGAGAG 721
CTCATCTTAAAGCTGAGTTCATCCCTGCGCTGAGGAGATCGGTCCGGGAGAG 983
CTCATCTTAAAGCTGAGTTCATCCCTGCGCTGAGGAGATCGGTCCGGGAGAG 781
CTCATCTTAAAGCTGAGTTCATCCCTGCGCTGAGGAGATCGGTCCGGGAGAG 1043
CTCATCTTAAAGCTGAGTTCATCCCTGCGCTGAGGAGATCGGTCCGGGAGAG 841
CTCATCTTAAAGCTGAGTTCATCCCTGCGCTGAGGAGATCGGTCCGGGAGAG 841
TGTGCTTCTGCAACAGTGGGCGGAGGCTCTCTCTCAACTTCAACTTCAACTG 1103
TGTGCTTCTGCAACAGTGGGCGGAGGCTCTCTCTCAACTTCAACTTCAACTG 901
TGTGCTTCTGCAACAGTGGGCGGAGGCTCTCTCTCAACTTCAACTTCAACTG 901
TGTGCTTCTGCAACAGTGGGCGGAGGCTCTCTCTCAACTTCAACTTCAACTG 1163
TGTGCTTCTGCAACAGTGGGCGGAGGCTCTCTCTCAACTTCAACTTCAACTG 961
TGTGCTTCTGCAACAGTGGGCGGAGGCTCTCTCTCAACTTCAACTTCAACTG 1223
TGTGCTTCTGCAACAGTGGGCGGAGGCTCTCTCTCAACTTCAACTTCAACTG 1021
TGTGCTTCTGCAACAGTGGGCGGAGGCTCTCTCTCAACTTCAACTTCAACTG 1283
TGTGCTTCTGCAACAGTGGGCGGAGGCTCTCTCTCAACTTCAACTTCAACTG 1081
TGTGCTTCTGCAACAGTGGGCGGAGGCTCTCTCTCAACTTCAACTTCAACTG 1343
TGTGCTTCTGCAACAGTGGGCGGAGGCTCTCTCTCAACTTCAACTTCAACTG 1343

Db 1082 CCAACATCCACAAATGAAAGCAATAAATCTACGTTGACTTGATGATAGAGCGAG 1141
 Qy 1344 CATGTACTCACCATGAGGAGCGCCCGGTCAAAAGAGCCGCAAGTTGTCCTGCGT 1403
 Db 1142 CATGTACTCACCATGAGGAGCGCCCGGTCAAAAGAGCCGCAAGTTGTCCTGCGT 1201
 Qy 1404 TTTCCGTGTCTGAATCTTCGACCTGCGATGACAACTTCACTGACATCTGGCTCCAA 1463
 Db 1202 TTTCCGTGTCTGAATCTTCGACCTGCGATGACAACTTCACTGACATCTGGCTCCAA 1261
 Qy 1464 ACACAAAATCTCTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1523
 Db 1262 ACACAAAATCTCTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1321
 Qy 1524 CATAGCTGACAGACACACGAGTACTGCAAGAAATCTTCACTGACGTTGCGCCAG 1583
 Db 1322 CATAGCTGACAGACACACGAGTACTGCAAGAAATCTTCACTGACGTTGCGCCAG 1381
 Qy 1584 TGACATCTTCACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1643
 Db 1382 TGACATCTTCACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1441
 Qy 1644 GAGCAGGCTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1703
 Db 1442 GAGCAGGCTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1501
 Qy 1704 GGCCTGCAACACAGCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1763
 Db 1502 GGCCTGCAACACAGCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1561
 Qy 1764 CTTCGCT 1823
 Db 1562 CTTCGCT 1621
 Qy 1824 GGTGACCTCTCGACCTTTCGACCTTTCGACCTTTCGACCTTTCGACCTTTCGACCTTTCGAC 1883
 Db 1622 GGTGACCTCTCGACCTTTCGACCTTTCGACCTTTCGACCTTTCGACCTTTCGACCTTTCGAC 1681
 Qy 1884 GGTGCTCTTATATCTTATATCTTATATCTTATATCTTATATCTTATATCTTATATCTTATAT 1943
 Db 1682 GGTGCTCTTATATCTTATATCTTATATCTTATATCTTATATCTTATATCTTATATCTTATAT 1741
 Qy 1944 AAGCAAGGTCTTACTGAGAGACCCCACTGAGAGACCCCACTGAGAGACCCCACTGAGAGACCC 2003
 Db 1742 AAGCAAGGTCTTACTGAGAGACCCCACTGAGAGACCCCACTGAGAGACCCCACTGAGAGACCC 1801
 Qy 2004 GTCTGGAACATCAGGCTGCGCAGAGACGAGTGGCTGCTGATCTTCTTAAAGAGCG 2063
 Db 1802 GTCTGGAACATCAGGCTGCGCAGAGACGAGTGGCTGCTGATCTTCTTAAAGAGCG 1861
 Qy 2064 GAGCGGCTGTGTCTGCGAGACAGGCGCGCATTCATGATCAGAGAGCAGCGAGCCG 2123
 Db 1862 GAGCGGCTGTGTCTGCGAGACAGGCGCGCATTCATGATCAGAGAGCAGCGAGCCG 1921
 Qy 2124 GGTGAGAGATCTTCAAGCTGAGACGAGATGTGTCTCCCAAGCAGCTTCCACATCA 2183
 Db 1922 GGTGAGAGATCTTCAAGCTGAGACGAGATGTGTCTCCCAAGCAGCTTCCACATCA 1981
 Qy 2184 CAGCTTGGGCTCAACATCTCTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2243
 Db 1982 CAGCTTGGGCTCAACATCTCTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2041
 Qy 2244 CTTCCTGGTGAACCTTACCCCAAG 2303
 Db 2042 CTTCCTGGTGAACCTTACCCCAAG 2101
 Qy 2304 GAGAGGTGAGATCTTACTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2360
 Db 2102 GAGAGGTGAGATCTTACTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2158

RESULT 6
 AAS21324

ID AAS21324 standard; cDNA, 1376 BP.
 XX
 AC AAS21324;
 XX
 DT 24-OCT-2001 (first entry)
 XX
 DE Human cDNA sequence encoding for PRO5773 polypeptide.
 XX
 KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;
 KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
 KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
 KW adipocyte; A-peptide; factor VIIA; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200140466-A2.
 XX
 PD 07-JUN-2001.
 XX
 PF 01-DEC-2000; 2000MO-US32678.
 XX
 PR 01-DEC-1999; 99MO-US28301.
 PR 01-DEC-1999; 99MO-US28634.
 PR 02-DEC-1999; 99MO-US28551.
 PR 02-DEC-1999; 99MO-US28564.
 PR 02-DEC-1999; 99MO-US28565.
 PR 09-DEC-1999; 99US-0170262.
 PR 16-DEC-1999; 99MO-US30095.
 PR 20-DEC-1999; 99MO-US30911.
 PR 20-DEC-1999; 99MO-US30999.
 PR 30-DEC-1999; 99MO-US31243.
 PR 06-JAN-2000; 2000MO-US00277.
 PR 06-JAN-2000; 2000MO-US00376.
 PR 11-FEB-2000; 2000MO-US03565.
 PR 18-FEB-2000; 2000MO-US04341.
 PR 18-FEB-2000; 2000MO-US04342.
 PR 22-FEB-2000; 2000MO-US04414.
 PR 24-FEB-2000; 2000MO-US04914.
 PR 24-FEB-2000; 2000MO-US05004.
 PR 01-MAR-2000; 2000MO-US05601.
 PR 20-MAR-2000; 2000MO-US07377.
 PR 21-MAR-2000; 2000MO-US07532.
 PR 30-MAR-2000; 2000MO-US08439.
 PR 17-MAY-2000; 2000MO-US13705.
 PR 22-MAY-2000; 2000MO-US14042.
 PR 30-MAY-2000; 2000MO-US14941.
 PR 02-JUN-2000; 2000MO-US15264.
 PR 10-NOV-2000; 2000MO-US30873.
 XX
 PA (GENTECH) GENENTECH INC.
 XX
 PI Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 XX
 DR WPI; 2001-408281/43.
 DR P-PSDB; AUI21252.
 XX
 PT Isolated, secretory and transmembrane PRO polypeptide used to detect
 PT other PRO polypeptides, link bioactive molecules to cells expressing
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
 PT lung, breast, prostate, cervical
 XX
 PS Claim 3; Fig 161; 813bp; English.
 XX
 CC AAS21244-AAS21518 encode for novel human secretory and transmembrane
 CC PRO polypeptides. The PRO polypeptides are useful to detect other
 CC PRO polypeptides, to link bioactive molecules to cells expressing
 CC PRO polypeptides, to modulate biological activities of cells expressing
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
 CC polypeptide expression in a cell sample to that in a control sample.
 CC Some of the 275 sequences are also useful to stimulate the release of

CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
CC proliferation or differentiation of chondrocytes, the proliferation or
CC gene expression in pericyte cells, the release of proteoglycans from
CC cartilage, the proliferation of inner ear utricular supporting cells or
CC of T-lymphocytes, the release of a cytokine from peripheral blood
CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of
CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
CC to factor VIIa. The PRO polypeptides can be used in assays to identify
CC molecules involved in binding interactions. The polynucleotides encoding
CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
CC transgenic or knock out animals and can be used in gene therapy.
XX
XX
SQ Sequence 1376 BP; 311 A; 384 C; 352 G; 329 T; 0 other;

Query Match 18.1%; Score 1117; DB 22; Length 1376;

Beet Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

190 CCAGGCGGAGGCGAGCTGCGCGGCTTGCGGCGCTGGGCGCGCGCTCCCGACCGTGC 249
1 CCAGGCGGAGGCGAGCTGCGCGGCTTGCGGCGCTGGGCGCGCGCTCCCGACCGTGC 60

250 TTTTCCCAACCGAGCGGAGCGGCTCCGAGTCATGCGCGGCTGAACTGCGGAGTCTCT 309
61 TTTTCCCAACCGAGCGGAGCGGCTCCGAGTCATGCGCGGCTGAACTGCGGAGTCTCT 120

310 ATCGACATGCTAGGGGTTTGTCTGCTGGGTGCGCGCGGCTCCCGCGCGGCGAGAGCT 369
121 ATCGACATGCTAGGGGTTTGTCTGCTGGGTGCGCGCGGCTCCCGCGCGGCGAGAGCT 180

370 TTGAGATTGCTCTGCGCAGAGAAAGCAATTACAGTTCTCTAATAAGCTGGGGAACCCG 429
181 TTGAGATTGCTCTGCGCAGAGAAAGCAATTACAGTTCTCTAATAAGCTGGGGAACCCG 240

430 ACTCTGCTGCAAAACCTGTTATCATGTCATTCTTAAAGACATATTAACAATGTTGTC 489
241 ACTCTGCTGCAAAACCTGTTATCATGTCATTCTTAAAGACATATTAACAATGTTGTC 300

490 ATCAAGCTGCGAGAAAGATAGTCTTACTTACTTACAGTCCAGAGTCTAGAGATCACTT 549
301 ATCAAGCTGCGAGAAAGATAGTCTTACTTACTTACAGTCCAGAGTCTAGAGATCACTT 360

550 GTCATAGAGATCCAGAAATAATTAATGATGATGTCAGGCGCATGTCCTTTTGGGAGGTT 609
361 GTCATAGAGATCCAGAAATAATTAATGATGATGTCAGGCGCATGTCCTTTTGGGAGGTT 420

610 CAGCTTCAGCCCTCGACATGTTGTTGCTTACCTCAACAGAACTTTCACTGGGATGTC 669
421 CAGCTTCAGCCCTCGACATGTTGTTGCTTACCTCAACAGAACTTTCACTGGGATGTC 480

670 AAAGCTATTAAGATCGGTTTAAAGCTGACGTTTCCATCCCTGCTGAGGCGATGTC 729
481 AAAGCTATTAAGATCGGTTTAAAGCTGACGTTTCCATCCCTGCTGAGGCGATGTC 540

730 GGTCCGGGAGAGAGCGCCAGAGCGAGTCATCTCATAGCGGCGGAAATCCATGTC 789
541 GGTCCGGGAGAGAGCGCCAGAGCGAGTCATCTCATAGCGGCGGAAATCCATGTC 600

790 ACCGTGCTCAGATCGGAACCTTCTGACAGCAATGGCACTGTGTCCGAGTCAAGATGCAA 849
601 ACCGTGCTCAGATCGGAACCTTCTGACAGCAATGGCACTGTGTCCGAGTCAAGATGCAA 660

850 GAAGAGTAAATAGGCTTAAACCTTCCATGAGTTTCAACCCAGAAATGTCCTCGGCTTC 909
661 GAAGAGTAAATAGGCTTAAACCTTCCATGAGTTTCAACCCAGAAATGTCCTCGGCTTC 720

910 AGCATTTGCAAAACCGCTCATCTATAAAGCTGTGTGATCATGAGAGTGTGTTTGAAGGT 969
721 AGCATTTGCAAAACCGCTCATCTATAAAGCTGTGTGATCATGAGAGTGTGTTTGAAGGT 780

970 GAAGGCTCAGCAACCTGTATGTCGCAACTAACAGAGGCTTCCCTGAGAGTGAAGTCTC 1029
|||||

DB 781 GAAGGCTCAGCAACCTGTATGTCGCAACTAACAGAGGCTTCCCTGAGATGAGTCTC 840
QY 1030 ATGAGCTGAGAGTTTGTGCTTCTGACACCTGAGGCGAGGCTCTCTTCTCAACTTC 1089
DB 841 ATGAGCTGAGAGTTTGTGCTTCTGACACCTGAGGCGAGGCTCTCTTCTCAACTTC 900
QY 1090 AACCTCTCCAACTGTAGAGAGAGAGAGGCGGTTGAATATCTACATCCCGGCTTCAAC 1149
DB 901 AACCTCTCCAACTGTAGAGAGAGAGAGGCGGTTGAATATCTACATCCCGGCTTCAAC 960
QY 1150 ACCAACCCCGAGAGTTTCAAGCTGAGAGCAACGAGCTGGGAAACATGCGGGAACCTTC 1209
DB 961 ACCAACCCCGAGAGTTTCAAGCTGAGAGCAACGAGCTGGGAAACATGCGGGAACCTTC 1020
QY 1210 AACCTCTCTGCAAGGCTGTGACCAAGTGCCTCAAGTCCGAGATCCTCGGCTGCGAG 1269
DB 1021 AACCTCTCTGCAAGGCTGTGACCAAGTGCCTCAAGTCCGAGATCCTCGGCTGCGAG 1080
QY 1270 TTCCAAGTTTGTGTCACATCCACAAATGAAAGCA 1306
DB 1081 TTCCAAGTTTGTGTCACATCCACAAATGAAAGCA 1117

RESULT 7
ID ACA03683
ACA03683 strand; cDNA; 1376 BP.

ACA03683;
23-MAY-2003 (first entry)

cDNA encoding human PRO polypeptide #81.

Human; PRO polypeptide; secreted and transmembrane protein;
tumour necrosis factor-alpha; TNF-alpha; blood; proliferation;
differentiation; chondrocyte; tumour; genetic disorder;
cytostatic; gene; ss.

OS Homo sapiens.

PN US2003036180-A1.

XX 20-FEB-2003.

PD 09-MAY-2002; 2002US-0143114.

PF 31-MAR-1997; 97WO-US05230.

XX 12-JUN-1998; 98WO-US12456.

PR 14-JUL-1998; 98WO-US14552.

PR 28-AUG-1998; 98WO-US17888.

PR 10-SEP-1998; 98WO-US18824.

PR 14-SEP-1998; 98WO-US19093.

PR 14-SEP-1998; 98WO-US19177.

PR 16-SEP-1998; 98WO-US19330.

PR 17-SEP-1998; 98WO-US19437.

PR 07-OCT-1998; 98WO-US21141.

PR 29-OCT-1998; 98WO-US22991.

PR 29-OCT-1998; 98WO-US22992.

PR 20-NOV-1998; 98WO-US24855.

PR 01-DEC-1998; 98WO-US25108.

PR 05-JAN-1999; 99WO-US00106.

PR 08-MAR-1999; 99WO-US05028.

PR 10-MAR-1999; 99WO-US05190.

PR 20-APR-1999; 99WO-US08615.

PR 14-MAY-1999; 99WO-US10733.

PR 02-JUN-1999; 99WO-US12252.

PR 01-SEP-1999; 99WO-US20111.

PR 08-SEP-1999; 99WO-US20594.

PR 13-SEP-1999; 99WO-US20944.

PR 15-SEP-1999; 99WO-US21090.

PR 05-OCT-1999; 99WO-US21547.

PR 99WO-US23089.

PR 29-NOV-1999; 99MO-US28214.
PR 30-NOV-1999; 99MO-US28313.
PR 30-NOV-1999; 99MO-US28409.
PR 01-DEC-1999; 99MO-US28301.
PR 01-DEC-1999; 99MO-US28634.
PR 02-DEC-1999; 99MO-US28551.
PR 02-DEC-1999; 99MO-US28564.
PR 02-DEC-1999; 99MO-US28565.
PR 16-DEC-1999; 99MO-US30095.
PR 20-DEC-1999; 99MO-US30911.
PR 20-DEC-1999; 99MO-US30999.
PR 22-DEC-1999; 99MO-US30720.
PR 30-DEC-1999; 99MO-US31243.
PR 30-DEC-1999; 99MO-US31274.
PR 05-JAN-2000; 2000MO-US00219.
PR 06-JAN-2000; 2000MO-US00277.
PR 06-JAN-2000; 2000MO-US00376.
PR 11-FEB-2000; 2000MO-US03565.
PR 18-FEB-2000; 2000MO-US04341.
PR 18-FEB-2000; 2000MO-US04342.
PR 22-FEB-2000; 2000MO-US04414.
PR 24-FEB-2000; 2000MO-US04914.
PR 24-FEB-2000; 2000MO-US05004.
PR 01-MAR-2000; 2000MO-US05601.
PR 02-MAR-2000; 2000MO-US05746.
PR 02-MAR-2000; 2000MO-US05841.
PR 10-MAR-2000; 2000MO-US06319.
PR 15-MAR-2000; 2000MO-US06884.
PR 20-MAR-2000; 2000MO-US07377.
PR 21-MAR-2000; 2000MO-US07532.
PR 30-MAR-2000; 2000MO-US08439.
PR 17-MAY-2000; 2000MO-US13705.
PR 22-MAY-2000; 2000MO-US14042.
PR 30-MAY-2000; 2000MO-US14941.
PR 02-JUN-2000; 2000MO-US15264.
PR 28-JUL-2000; 2000MO-US20710.
PR 11-AUG-2000; 2000MO-US22031.
PR 23-AUG-2000; 2000MO-US23522.
PR 24-AUG-2000; 2000MO-US23328.
PR 08-NOV-2000; 2000MO-US30952.
PR 10-NOV-2000; 2000MO-US30873.
PR 01-DEC-2000; 2000MO-US32678.
PR 20-DEC-2000; 2000MO-US34956.
PR 28-FEB-2001; 2001MO-US06520.
PR 01-MAR-2001; 2001MO-US06666.
PR 25-MAY-2001; 2001MO-US17092.
PR 01-JUN-2001; 2001MO-US17800.
PR 20-JUN-2001; 2001MO-US19692.
PR 22-JUN-2001; 2001MO-US20116.
PR 29-JUN-2001; 2001MO-US21066.
PR 09-JUL-2001; 2001MO-US21735.
PR 20-DEC-2000; 2000MO-US27259.
PR 28-FEB-2001; 2001MO-US27498.
PR 09-MAR-2001; 2001MO-US082706.
PR 14-MAR-2001; 2001MO-US08689.
PR 22-MAR-2001; 2001MO-US086744.
PR 05-APR-2001; 2001MO-US08366.
PR 10-MAY-2001; 2001MO-US084208.
PR 18-MAY-2001; 2001MO-US084280.
PR 10-MAY-2001; 2001MO-US0854280.
PR 25-MAY-2001; 2001MO-US0860216.
PR 25-MAY-2001; 2001MO-US086028.
PR 01-JUN-2001; 2001MO-US0866034.
PR 05-JUN-2001; 2001MO-US0872035.
PR 14-JUN-2001; 2001MO-US0874503.
PR 19-JUN-2001; 2001MO-US0882636.
PR 21-JUN-2001; 2001MO-US0886342.
PR 18-JUL-2001; 2001MO-US0887879.
PR 06-AUG-2001; 2001MO-US0924419.
PR 09-AUG-2001; 2001MO-US0927796.
PR 16-AUG-2001; 2001MO-US0931836.
PR 19-DEC-2001; 2001MO-US0028072.
XX

PA (GETH) GENENTECH INC.
XX Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;
PI Geritsen ME, Goddard A, Godowski PJ, Gunney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX WPI, 2003-332040/31.
DR P-PSDB; AB066650.
XX
XX New secreted and transmembrane PRO nucleic acids, useful for gene
PT therapy, in chromosome and gene mapping, as chromosome markers, in
PT tissue typing, and in chromosome identification -
XX
XX Claim 2; Fig 161; 660pp; English.
PS
XX
CC The present invention relates to the isolation of novel human PRO
CC polypeptides, and the polynucleotide sequences encoding them. The
CC PRO polypeptides are secreted and transmembrane proteins. The PRO
CC polypeptides are useful for detecting other PRO polypeptides, for
CC linking bioactive molecules to cells expressing PRO polypeptides,
CC for modulating biological activities of cells expressing PRO
CC polypeptides, and for identifying agonists or antagonists.
CC The PRO polypeptides are useful for stimulating the release of
CC tumour necrosis factor (TNF)-alpha from human blood, for stimulating the
CC proliferation or differentiation of chondrocytes, and detecting the
CC presence of tumours. The polynucleotide sequences encoding PRO
CC polypeptides are useful as hybridisation probes, in chromosome and
CC gene mapping, in the generation of antisense RNA and DNA, in the
CC preparation of PRO polypeptides, for generating transgenic animals or
CC knockout animals, for the genetic analysis of individuals with genetic
CC disorders, and in gene therapy. ACA03603-ACA03877 represent cDNAs
CC encoding the human PRO polypeptides of the invention.
CC Note: The sequence data for this patent was obtained in electronic
CC format directly from the USPTO web site at
CC segdata.uspto.gov/psipdsidentry.html.
XX
SQ Sequence 1376 BP; 311 A; 384 C; 352 G; 329 T; 0 other;

Query Match 18.1%; Score 1117; DB 25; Length 1376;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 CCAAGGCGGAGCGGAGCTGCGCGGCTTGGGCGCTGGGCGCTCCACCGCTG 249
DB 1 CCAGGGCGGAGCGGAGCTGCGCGGCTTGGGCGCTGGGCGCTCCACCGCTG 60
QY 250 TTTTCCCGGAGCGGCGGCTGCGCGGAGCTGCGCGGCTGGAAGTGGGGGCTCT 309
DB 61 TTTTCCCGGAGCGGCGGCTGCGCGGAGCTGCGCGGCTGGAAGTGGGGGCTCT 120
QY 310 ATCGCACTGCTAGGGGTTCTGCTGCTGCGGCGGCGCTGCGGCGGCGGAGAGCT 369
DB 121 ATCGCACTGCTAGGGGTTCTGCTGCTGCGGCGGCGGCGGCGGCGGAGAGCT 180
QY 370 TTTGAGATTGCTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 429
DB 181 TTTGAGATTGCTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
QY 430 ACTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 489
DB 241 ACTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
QY 490 ATCAAGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 549
DB 301 ATCAAGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
QY 550 GTCATAGAGATCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 609
DB 361 GTCATAGAGATCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
QY 610 CAGCTTCAGGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 669
DB 421 CAGCTTCAGGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480

29-JUN-2001, 2001WO-US21066.
PR 09-JUL-2001, 2001WO-US21735.
PR 20-DEC-2000, 2000US-0747259.
PR 28-FEB-2001, 2001US-0796498.
PR 09-MAR-2001, 2001US-0802706.
PR 14-MAR-2001, 2001US-0808689.
PR 22-MAR-2001, 2001US-0816744.
PR 05-APR-2001, 2001US-0828366.
PR 10-MAY-2001, 2001US-0854208.
PR 10-MAY-2001, 2001US-0854280.
PR 18-MAY-2001, 2001US-0860216.
PR 25-MAY-2001, 2001US-0866028.
PR 25-MAY-2001, 2001US-0866034.
PR 01-JUN-2001, 2001US-0872035.
PR 05-JUN-2001, 2001US-0874503.
PR 14-JUN-2001, 2001US-0882636.
PR 19-JUN-2001, 2001US-0886342.
PR 21-JUN-2001, 2001US-0887879.
PR 18-JUL-2001, 2001US-0908827.
PR 06-AUG-2001, 2001US-0924419.
PR 09-AUG-2001, 2001US-0927796.
PR 16-AUG-2001, 2001US-0931836.
PR 19-DEC-2001, 2001US-0028072.

(GERTH) GENENTECH INC.

Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W,
PI Geritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX MPI, 2003-331925/31.
DR P-PSDB; ABUS6926.

PT New secreted and transmembrane nucleic acids and polypeptides,
PT designated as PRO, useful for treating inflammation, organ failure,
PT atherosclerosis, cardiac injury, infertility, birth defects, premature
PT aging, AIDS, or cancer -

Claim 2; Fig 161; 659pp; English.

XX The invention relates to an isolated nucleic acid comprising, or which is
XX at least 80% identical to, or the full-length coding sequence of, any of
XX the 275 nucleotide sequences, encoding the corresponding PRO polypeptide
XX (one of 275 secreted or transmembrane proteins). The nucleic acid
XX further comprises the full-length coding sequence of the DNA deposited
XX under American Type Culture Collection (ATCC) accession number in a list
XX given in the specification. Also included are vectors and host
XX cells for producing PRO proteins, PRO fusion proteins, anti-PRO
XX antibodies, PRO extracellular domains and mature sequences, methods
XX of detecting PRO proteins, methods for stimulating the release of
XX TNF-alpha (tumour necrosis factor alpha) from human blood,
XX (and the proliferation of differentiation of chondrocyte cells, the
XX proliferation of, or gene expression in pericyte cells, the release or
XX proteoglycans from cartilage, proliferation of inner ear utricular
XX supporting cells, the proliferation of T-lymphocyte cells, the release
XX of a cytokine from peripheral blood mononuclear cells (PBMC), or the
XX proliferation of endothelial cells), a method for modulating the uptake
XX of glucose or free fatty acid (FFA) by skeletal muscle cells,
XX a method for inhibiting the binding of A-peptide to factor VIIA,
XX or the differentiation of adipocyte cells, a method for detecting the
XX presence of a tumour in a mammal and an oligonucleotide probe derived
XX from any of the nucleotide sequences cited above. The nucleic acids and
XX polypeptides are useful for treating inflammatory diseases, organ
XX failure, atherosclerosis, cardiac injury, infertility, birth defects,
XX premature aging, AIDS (acquired immunodeficiency syndrome), cancer, or
XX diabetic complications. The nucleic acids are useful as hybridisation
XX probes, in chromosome and gene mapping, and in generating antisense RNA
XX or DNA. The polypeptides are useful as pharmaceuticals, diagnostics,
XX biosensors or bioreactors. Both are useful in tissue typing.
XX The present sequence encodes a PRO protein of the invention.

Sequence 1376 BP; 311 A; 384 C; 352 G; 329 T; 0 other;

Query Match	18.1%	Score 1117	DB 25	Length 1376
Best Local Similarity	100.0%	Pred. No. 0	Mismatches 0	Indels 0
Matches 1117	Conservative 0	Indels 0	Gaps 0	
QY 190	CCAGGAGCCGAGCGAGCTGCGCCGCGGCTTGGGCGCTGGGCGCGCGCTCCCGACGCTG	249		
Db 1	CCAGGAGCCGAGCGAGCTGCGCCGCGGCTTGGGCGCTGGGCGCGCGCTCCCGACGCTG	60		
QY 250	TTTTTCCCAACCGAGCCGAGCGCTCCGAGTCAATGCGCGGCTGAACTGCGGGCTCTT	309		
Db 61	TTTTTCCCAACCGAGCCGAGCGCTCCGAGTCAATGCGCGGCTGAACTGCGGGCTCTT	120		
QY 310	ATCGCACTGCTAGGGGCTTCTGCTGGTGGGCGCGCGCTGCGCGGCGGAGAGCT	369		
Db 121	ATCGCACTGCTAGGGGCTTCTGCTGGTGGGCGCGCGCTGCGCGGCGGAGAGCT	180		
QY 370	TTTGAGATTGCTCTGCGCGAGAGAGCAATTAAGTTCTGATAAGCTGGGAGACCCG	429		
Db 181	TTTGAGATTGCTCTGCGCGAGAGAGCAATTAAGTTCTGATAAGCTGGGAGACCCG	240		
QY 430	ACTGTGCTGGCAAAACCGTGTACATGCTATTTCTAAAGACATATAACATGTTGTCC	489		
Db 241	ACTGTGCTGGCAAAACCGTGTACATGCTATTTCTAAAGACATATAACATGTTGTCC	300		
QY 490	ATCAAGCTCGAGAAAGATAGCTTTTACCTTACCTGCGAGAGTCTGAGATCACTT	549		
Db 301	ATCAAGCTCGAGAAAGATAGCTTTTACCTTACCTGCGAGAGTCTGAGATCACTT	360		
QY 550	GTCATAGAGATCGAGAAATATTGACTGTATGAGGCGCATGCTTTGGGAGGTT	609		
Db 361	GTCATAGAGATCGAGAAATATTGACTGTATGAGGCGCATGCTTTGGGAGGTT	420		
QY 610	CAGCTTACGCTCGACATGCTTGTGCTTACCTTCAACAGAACTTCACTGGAGATGTC	669		
Db 421	CAGCTTACGCTCGACATGCTTGTGCTTACCTTCAACAGAACTTCACTGGAGATGTC	480		
QY 670	AAAGCTCATAGAGATGCTTTTAAAGCTGACGTTTTCATCCCTGCTGAGGAGATC	729		
Db 481	AAAGCTCATAGAGATGCTTTTAAAGCTGACGTTTTCATCCCTGCTGAGGAGATC	540		
QY 730	GGTCCGGGGGAGAGGCTGCGCAGAGAGTCACTCATGAGGCGCGGAGTCAATGTC	789		
Db 541	GGTCCGGGGGAGAGGCTGCGCAGAGAGTCACTCATGAGGCGCGGAGTCAATGTC	600		
QY 790	ACCGTGTGAGAGTGAAGACCTTCTGAGCAATGAGCTGTGTCCCGATCAAGATGCA	849		
Db 601	ACCGTGTGAGAGTGAAGACCTTCTGAGCAATGAGCTGTGTCCCGATCAAGATGCA	660		
QY 850	GAGAGTGAAGAAATGCGCTTACCTCCCATGCTTCCAGAAATGTCTCCGCTTC	909		
Db 661	GAGAGTGAAGAAATGCGCTTACCTCCCATGCTTCCAGAAATGTCTCCGCTTC	720		
QY 910	ACCATTTGCAAAACCGCTCATCTTAAAGCTGTGACATCATGAGTGTGTTGAGGT	969		
Db 721	ACCATTTGCAAAACCGCTCATCTTAAAGCTGTGACATCATGAGTGTGTTGAGGT	780		
QY 970	GAAAGCTGAGCAACCTGATGCTGCAACCTACCCAGAGAGCTTCCGTGAGATGAGCTC	1029		
Db 781	GAAAGCTGAGCAACCTGATGCTGCAACCTACCCAGAGAGCTTCCGTGAGATGAGCTC	840		
QY 1030	ATGAGTGGCAGTGTGTTGCTTCTGCAACCTGCGGCGCGAGGTCTCTTCTCAACTTC	1089		
Db 841	ATGAGTGGCAGTGTGTTGCTTCTGCAACCTGCGGCGCGAGGTCTCTTCTCAACTTC	900		
QY 1090	AACCTCTCAACTGTGAGAGAGAGAGAGAGGAGGTTGAATCTACATCCCGGCTCCAC	1149		
Db 901	AACCTCTCAACTGTGAGAGAGAGAGAGGAGGTTGAATCTACATCCCGGCTCCAC	960		
QY 1150	ACCAACCCGAGAGTGTTCAGCTGAGAGCAAGCAAGCTTGGAGAAATGGCGGAACTTC	1209		
Db 961	ACCAACCCGAGAGTGTTCAGCTGAGAGCAAGCAAGCTTGGAGAAATGGCGGAACTTC	1020		
QY 1210	AACCTCTCTGCAAGGCTGTGACCAAGATGCCAAAGTCCAGGAGTCTCCGCGCTGAG	1269		

Db 1021 AACCTCTCTCTCAAGGCTGTGACCAAGTGCACCAAGTCTCTCGGCTGCAG 1080
Oy 1270 TTCCAAGTTTGGTCCACATCCACAAATGAAGCA 1306
Db 1081 TTCCAAGTTTGGTCCACATCCACAAATGAAGCA 1117

RESULT 9
ABX89221
ID ABX89221 standard; cDNA; 1376 BP.
XX
AC ABX89221;
XX
DT 13-MAY-2003 (first entry)
XX
DE DNA encoding novel secreted and transmembrane protein PRO5773.
XX
KW Human; PRO; hypertrophy of neonatal heart; angiogenesis; wound healing;
KW cardiac insufficiency disorder; cancer; tumour; immune response;
KW adrenal cortical capillary endothelial growth; c-fos induction;
KW vascular endothelial growth factor inhibition; VEGF inhibition;
KW endothelial cell growth inhibitor; T-lymphocytes stimulation;
KW retinal neurons cell survival; rod photoreceptor cell survival;
KW retinal disorder; retinitis pigmentosa; kidney disorder;
KW mammalian kidney mesangial cell proliferation; Berger disease;
KW dermatitis; herpeticiformis; Crohn's disease; chondrocyte proliferation;
KW chondrocyte redifferentiation; sports injury; arthritis; gene; ss.
XX
OS Homo sapiens.
XX
PN US2003017563-A1.
XX
PD 23-JAN-2003.
XX
PF 07-MAY-2002; 2002US-0140808.
XX
XX 31-MAR-1997; 97WO-US05230.
PR 12-JUN-1998; 98WO-US12456.
PR 14-JUL-1998; 98WO-US14552.
PR 28-AUG-1998; 98WO-US17888.
PR 10-SEP-1998; 98WO-US18824.
PR 14-SEP-1998; 98WO-US19093.
PR 14-SEP-1998; 98WO-US19094.
PR 14-SEP-1998; 98WO-US19177.
PR 16-SEP-1998; 98WO-US19330.
PR 17-SEP-1998; 98WO-US19437.
PR 07-OCT-1998; 98WO-US21411.
PR 29-OCT-1998; 98WO-US22991.
PR 29-OCT-1998; 98WO-US22992.
PR 20-NOV-1998; 98WO-US24855.
PR 01-DEC-1998; 98WO-US25108.
PR 05-JAN-1999; 99WO-US00106.
PR 08-MAR-1999; 99WO-US05028.
PR 10-MAR-1999; 99WO-US05190.
PR 20-APR-1999; 99WO-US08615.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 30-NOV-1999; 99WO-US28409.
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.

PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 22-DEC-1999; 99WO-US30720.
PR 30-DEC-1999; 99WO-US31243.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00277.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05746.
PR 02-MAR-2000; 2000WO-US05841.
PR 10-MAR-2000; 2000WO-US06319.
PR 15-MAR-2000; 2000WO-US06884.
PR 20-MAR-2000; 2000WO-US07377.
PR 21-MAR-2000; 2000WO-US07532.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US13705.
PR 30-MAY-2000; 2000WO-US14042.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUN-2000; 2000WO-US20710.
PR 11-AUG-2000; 2000WO-US22031.
PR 23-AUG-2000; 2000WO-US23522.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 10-NOV-2000; 2000WO-US30873.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-MAR-2001; 2001WO-US06666.
PR 25-MAY-2001; 2001WO-US17092.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 22-JUN-2001; 2001WO-US20116.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 20-DEC-2000; 2000US-0742259.
PR 28-FEB-2001; 2001US-0796498.
PR 09-MAR-2001; 2001US-0803706.
PR 14-MAR-2001; 2001US-0806889.
PR 22-MAR-2001; 2001US-0816744.
PR 05-APR-2001; 2001US-0828366.
PR 10-MAY-2001; 2001US-0854208.
PR 10-MAY-2001; 2001US-0854280.
PR 18-MAY-2001; 2001US-0860216.
PR 25-MAY-2001; 2001US-086028.
PR 05-MAY-2001; 2001US-086034.
PR 01-JUN-2001; 2001US-0872035.
PR 05-JUN-2001; 2001US-0874503.
PR 14-JUN-2001; 2001US-0882636.
PR 19-JUN-2001; 2001US-0886342.
PR 21-JUN-2001; 2001US-0887879.
PR 18-JUL-2001; 2001US-0908827.
PR 06-AUG-2001; 2001US-0924419.
PR 09-AUG-2001; 2001US-0927796.
PR 16-AUG-2001; 2001US-0931836.
PR 19-DEC-2001; 2001US-0028072.
XX
XX (GENTH) GENENTECH INC.
XX
XX Baker KP, Beresini M, Deforge L, Desnoyers L, Pilyaroff B, Gao W;
PI Gertsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WT, Zhang Z;
XX WPI; 2003-148238/14.
DR P-PsDB; ABUS9731.
XX

PR 13-DEC-2000; 2000US-255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer; stage of prostate cancer -
XX
XX Claim 1; Page 6580; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 1031 BP; 272 A; 255 C; 212 G; 263 T; 29 other;
XX
XX
XX Query Match 10.9%; Score 669; DB 23; Length 1031;
XX Best Local Similarity 100.0%; Pred. NO. 1.9e-293;
XX Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 4499 CTTTTCCTTTTCTTTTATGAGACAGATCTCTGTCCTAGGCTGAGTGC 4558
XX 30 CTTTTCCTTTTCTTTTATGAGACAGATCTCTGTCCTAGGCTGAGTGC 89
XX
XX 4559 AGTGGCGCAATCTGGCTCACTGCAACCTCTGCTCTGAGCTCAAGCAATCTCCAC 4618
XX 90 AGTGGCGCAATCTGGCTCACTGCAACCTCTGCTCTGAGCTCAAGCAATCTCCAC 149
XX
XX 4619 TGAGCTCCCAATAGCTGGATCACTGGCACAACACCACTGCTAATTTTGTAT 4678
XX 150 TGAGCTCCCAATAGCTGGATCACTGGCACAACACCACTGCTAATTTTGTAT 209
XX
XX 4679 TTTTGTAGAGACAGGGTTTCAACATGTTGCCAGGCTGATCTCAACTCTGAGCTCAA 4738
XX 210 TTTTGTAGAGACAGGGTTTCAACATGTTGCCAGGCTGATCTCAACTCTGAGCTCAA 269
XX
XX 4739 GCAATCTCTGCTGGCTGCGCTCCCAAAAGTGGGATTAAGATGTAGCCACCGCATCC 4798
XX 270 GCAATCTCTGCTGGCTGCGCTCCCAAAAGTGGGATTAAGATGTAGCCACCGCATCC 329
XX
XX 4799 ACCCCACACCCCTCATTTTATACCAATTACTGCGCCAGTAACTGTGACTTTTCTTCC 4858
XX 330 ACCCCACACCCCTCATTTTATACCAATTACTGCGCCAGTAACTGTGACTTTTCTTCC 389
XX
XX 4859 ACCCCTGCTGTAGTGAAGAGAGAGATATGATATGCTTGTAGAGACAGTCCCAAG 4918
XX 390 ACCCCTGCTGTAGTGAAGAGAGAGATATGATATGCTTGTAGAGACAGTCCCAAG 449
XX
XX 4919 TTCAATATTTTCTGCGGCAAAAATCTCTTCAAAAATAATGTAATCTTATTTCAAT 4978
XX 450 TTCAATATTTTCTGCGGCAAAAATCTCTTCAAAAATAATGTAATCTTATTTCAAT 509
XX
XX 4979 GAATTCACCTTGAATGACACCGCTCACTTGTTCATCATGATGAATGAAGAAATTT 5038
XX 510 GAATTCACCTTGAATGACACCGCTCACTTGTTCATCATGATGAATGAAGAAATTT 569
XX
XX 5039 TATAGTCTCTTAAATGGCGTGTACTGCAAGACCTTTTGAACACTTTCAGAGAGATGAT 5098

DB 570 TATAGTCTCTTAAATGGCGTGTACTGCAAGACCTTGAACACTTTCAGAGAGATGAT 629
XX
XX 5099 ATTTAAGTCATGCGCTTGGCGCTGCTTATGACACTTCCCTTGAAGTGTGTTCT 5158
XX 630 ATTTAAGTCATGCGCTTGGCGCTGCTTATGACACTTCCCTTGAAGTGTGTTCT 689
XX
XX 5159 GCCCAGTGA 5167
XX 690 GCCCAGTGA 698
XX
XX
XX RESULT 11
XX AAA78075
XX ID AAA78075 standard; cDNA; 544 BP.
XX
XX AAA78075;
XX
XX 14-NOV-2000 (first entry)
XX
XX cDNA encoding human colon tumour polypeptide, SEQ ID NO:362.
XX
XX Human colon tumour polypeptide; tumour antigen; cancer; vaccine;
XX immunotherapy; diagnosis; progression; ss.
XX
XX Homo sapiens.
XX
XX WO200037643-A2.
XX
XX 29-JUN-2000.
XX
XX 23-DEC-1999; 99WO-US30909.
XX
XX 23-DEC-1998; 98US-0221298.
XX 02-JUL-1999; 99US-0347496.
XX 22-SEP-1999; 99US-0401064.
XX 19-NOV-1999; 99US-0444242.
XX 02-DEC-1999; 99US-0454150.
XX
XX (CORI-) CORIXA CORP.
XX
XX Xu J, Lodes MJ, Secretat H, Benson DR, Meagher MJ, Stolk J;
XX Wang T, Yugu J;
XX
XX WPI; 2000-442671/38.
XX
XX New colon tumor polypeptides used to inhibit the development of cancer,
XX especially colon cancer, and for diagnosing and monitoring the
XX progression of the cancer -
XX
XX Claim 1; Page 196; 229pp; English.
XX
XX Sequences AAA77722-A78199 represent 478 cDNAs encoding proteins or
XX portions of proteins which are associated with human colon tumours.
XX The invention also specifically discloses 8 human colon tumour proteins
XX (AAH1897-811904). The nucleic acids, the polypeptides they encode, and
XX antigen presenting cells (APCs; preferably dendritic cells) expressing
XX such polypeptides may be used in vaccines that target tumour cells,
XX especially colon tumour cells, thereby inhibiting the development of
XX cancer. T-cells specific for the polypeptide expressed by the APC are
XX used to remove tumour cells from biological samples, especially blood or
XX fractions thereof. The sample or the isolated T-cells specific for the
XX polypeptide can then be used to inhibit cancer development. CD4+ and/or
XX CD8+ T-cells from a patient may be incubated with a polypeptide or
XX nucleic acid of the invention, or an APC expressing such a polypeptide,
XX to cause the proliferation of specific T-cells. The T-cells can be
XX cloned and then administered back to the patient to inhibit cancer
XX development. Nucleic acids encoding the polypeptides and antibodies
XX against the polypeptides may be used to determine the expression level
XX of a tumour protein of the invention, and therefore to determine whether
XX cancer cells are present. Such diagnostic methods may also be used to
XX monitor the progression of a cancer by repeating the processes at time
XX intervals, and comparing the current result to previous results. The

Qy	3428	TGTTTGAGCAGCATTTGACACATATCTGTTTGATTAAGACCTTCTGATTCCTAGCTG	3487
Db	361	TGTTTGACAGCAGCATTTGACACATATCTGTTTGATTAAGACCTTCTGATTCCTAGCTG	420
Qy	3488	GTTCGTGGTATATCCCATTTGTGAAATTCATCTTGATTCATCTGCTATAGTCTAGCA	3547
Db	421	GTTCGTGGTATATCCCATTTGTGAAATTCATCTTGATTCATCTGCTATAGTCTAGCA	480
Qy	3548	ATAAGAGAAATTTTCTCTCAAGTTTTCATGTGGCGTTCTCTAGCTGACCAATACCTTGAC	3607
Db	481	ATAAGAGAAATTTTCTCTCAAGTTTTCATGTGGCGTTCTCTAGCTGACCAATACCTTGAC	540
Qy	3608	ATTT 3611	
Db	541	ATTT 544	
RESULT 13			
ID	AB232999	standard; cDNA; 544 BP.	
XX	AB232999;		
XX	30-JAN-2003	(first entry)	
DE	Human colon tumour CDNA clone 25928 SEQ ID NO:362.		
KX	Human; colon cancer; colon tumour; immunotherapy; diagnosis; cancer;		
KW	tumour; immune response; immunostimulant; cytostatic; vaccine;		
XX	gene; ss.		
OS	Homo sapiens.		
XX	WO200283070-A2.		
XX	24-OCT-2002.		
XX	09-APR-2002; 2002WO-US11475.		
XX	10-APR-2001; 2001US-0833263.		
PR	03-AUG-2001; 2001US-0922217.		
PR	19-DEC-2001; 2001US-0025380.		
XX	(CORI-) CORIXA CORP.		
PI	Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;		
PI	Wang T, Jiang Y, Smith CL, King GE, Wang A, Clapper JD;		
PI	Skelsky YAM, Fanger GR, Vedvick TS, Carter D;		
XX	WPI; 2003-067548/06.		
DR	New polynucleotide, useful for the preparation of a composition for		
PT	stimulating an immune response against, or treating, cancer -		
XX	Example 1; Page 236-237; 537pp; English.		
PS	The present invention describes compounds (I) for the immunotherapy and		
XX	diagnosis of colon cancer. Also described: (1) a method for detecting		
CC	the presence of cancer in a patient; (2) a method for stimulating and/or		
CC	expanding T cells specific for a tumour protein; (3) an isolated T cell		
CC	population comprising T cells prepared by the method of (2); (4) a method		
CC	for stimulating an immune response in a patient; (5) a method for		
CC	treating cancer in a patient; and (6) a method for inhibiting the		
CC	cytostatic activities and can be used in vaccines. AB232646 to AB233725		
CC	and AB235343 to AB235351 represent human colon cancer/tumour related		
CC	sequences used in the exemplification of the present invention.		
XX	Sequence 544 BP; 138 A; 123 C; 113 G; 170 T; 0 other;		
Qy	Query Match	8.0%; Score 493; DB 25; Length 544;	
XX	Best Local Similarity	99.8%; Pred. No. 1.9e-21;	
XX	Matches 543; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		

Chr	Start (Mb)	End (Mb)	Gene	Strand	RefSeq ID	RefSeq Description	RefSeq Length (bp)
1	100,000,000	100,000,500	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,000,500	100,001,000	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,001,000	100,001,500	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,001,500	100,002,000	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,002,000	100,002,500	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,002,500	100,003,000	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,003,000	100,003,500	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,003,500	100,004,000	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,004,000	100,004,500	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,004,500	100,005,000	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,005,000	100,005,500	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,005,500	100,006,000	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,006,000	100,006,500	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,006,500	100,007,000	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,007,000	100,007,500	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,007,500	100,008,000	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,008,000	100,008,500	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,008,500	100,009,000	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,009,000	100,009,500	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,009,500	100,010,000	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,010,000	100,010,500	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,010,500	100,011,000	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,011,000	100,011,500	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,011,500	100,012,000	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,012,000	100,012,500	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,012,500	100,013,000	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,013,000	100,013,500	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,013,500	100,014,000	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,014,000	100,014,500	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,014,500	100,015,000	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,015,000	100,015,500	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,015,500	100,016,000	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,016,000	100,016,500	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,016,500	100,017,000	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,017,000	100,017,500	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,017,500	100,018,000	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1	100,018,000	100,018,500	ABT2161	+	ABT2161	ABT2161 standard, DNA; 660 BP.	660
1							

PT Novel isolated polypeptide encoded by breast cancer marker gene, useful
PT for diagnosing, staging, monitoring, prognosing and treating diseases
PT associated with breast cancer -
XX
XX
PS Disclosure, Page 101; 725pp; English.
XX
CC The invention relates to an isolated polypeptide encoded by a breast
CC cancer marker gene comprising any of 1417 21-805 nucleotide sequences,
CC given in the specification. The methods of the invention are useful for
CC diagnosing patients having an identified breast mass or symptoms
CC associated with breast cancer, to diagnose breast cancer or its
CC precursors, and for monitoring the efficacy of treatment of a breast
CC cancer patient (e.g. efficacy of chemotherapy). The methods are also
CC useful for evaluating a patient before, after or during therapy, to
CC evaluate the reduction in a tumour burden. The breast cancer marker gene
CC proteins are useful as immunogens for raising antibodies, by immunising a
CC mammal with a breast cancer marker protein. The marker proteins are
CC useful as bait proteins in a two-hybrid or three-hybrid assay, to
CC identify other proteins which bind to or interact with the marker
CC proteins. The breast cancer marker genes are useful as surrogate marker
CC genes for one or more disorders, disease states or conditions leading to
CC disease states, in particular, breast cancers. The breast cancer marker
CC genes are useful as pharmacodynamic marker genes. An antibody which
CC selectively binds to a protein of a breast cancer marker gene is useful
CC for treating cancers, particularly breast cancers. The host cell of the
CC invention is useful for producing non-human transgenic animals. This
CC polynucleotide sequence represents one of the breast cancer marker genes
CC of the invention.
XX
XX

Sequence 660 BP; 162 A; 159 C; 182 G; 154 T; 3 other;

Query Match 7.9%; Score 486; DB 25; Length 660;
Best Local Similarity 100.0%; Pred. No. 2.9e-210;
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5421 CCACGCGCCATTTTCTTAACTGAGGGCTCAAACTCCTGCAAGTGTGCTCCT 5480
DB 532 CCACGCGCCATTTTCTTAACTGAGGGCTCAAACTCCTGCAAGTGTGCTCCT 473
QY 5481 GAGACCAAGTATTCCTGAGCTGAGCTCAGTAAAGGGGCCAGCCTGAGAAACCTGGC 5540
DB 472 GAGACCAAGTATTCCTGAGCTGAGCTCAGTAAAGGGGCCAGCCTGAGAAACCTGGC 413
QY 5541 TCTTTTCTTAAAGCCAGGCCACCTTACATTAACATTTTCAAGGTCATCTGAAACAGT 5600
DB 412 TCTTTTCTTAAAGCCAGGCCACCTTACATTAACATTTTCAAGGTCATCTGAAACAGT 353
QY 5601 GAAGTGCATTTTGTGAAGCTTCTGATCCAGCCCATCTGCTCATCTGATGCTGCC 5660
DB 352 GAAGTGCATTTTGTGAAGCTTCTGATCCAGCCCATCTGCTCATCTGATGCTGCC 293
QY 5661 ATCCCTACGAGAAAGCCAGCATGAGAGCTGCTCTTAATGCTGTGTCATTCACA 5720
DB 292 ATCCCTACGAGAAAGCCAGCATGAGAGCTGCTCTTAATGCTGTGTCATTCACA 233
QY 5721 GAAGGAAAGGTCTCAAGAAAGTCAACTGAGGACAAGCAAGCCACCGACATGGCC 5780
DB 232 GAAGGAAAGGTCTCAAGAAAGTCAACTGAGGACAAGCAAGCCACCGACATGGCC 173
QY 5781 TTGGTAAAGGTTCAGACATGCTGTGTGTGATCTGCAAGTCTTCACTGGAATTAATTTA 5840
DB 172 TTGGTAAAGGTTCAGACATGCTGTGTGTGATCTGCAAGTCTTCACTGGAATTAATTTA 113
QY 5841 TTCAATGAGATCTTTTATGAGTGGCATTTTATTCATTTCTGTGCTTTAAATTAACAA 5900
DB 112 TTCAATGAGATCTTTTATGAGTGGCATTTTATTCATTTCTGTGCTTTAAATTAACAA 53
QY 5901 TGTACC 5906
DB 52 TGTACC 47

RESULT 15

AAS64483
ID AAS64483 standard; CDNA; 487 BP.
XX
XX AAS64483;
AC
XX
DT 13-FEB-2002 (first entry)
XX
XX
DE DNA encoding novel human diagnostic protein #287.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Dormanac RT, Liu C, Tang YT;
XX
XX P-PSDB; ABG00296.
XX
XX WPI; 2001-639362/73.
XX
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -
XX
XX
XX Claim 1, SEQ ID No 287; 103pp; English.
XX
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human
XX CC diagnostic coding sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 487 BP; 131 A; 145 C; 120 G; 91 T; 0 other;

Query Match 7.8%; Score 482; DB 23; Length 487;
Best Local Similarity 100.0%; Pred. No. 1.9e-208;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2312 GAGCTTACTGCTGTGCTGCGCTCGGCTCATATTTGCTGTGTGAAAAAGAAAAA 2371
DB 6 GAGCTTACTGCTGTGCTGCGCTCGGCTCATATTTGCTGTGTGAAAAAGAAAAA 65
QY 2372 AGACAAACAAAGGCGCCGCTGAGGATCTACAATGAGCAACATCACTAGATGCGGA 2431
DB 66 AGACAAACAAAGGCGCCGCTGAGGATCTACAATGAGCAACATCACTAGATGCGGA 125

```
OY 2432 GGCAGCCAAAAAGTTTCAGAAAGGCGGAAAGACATGACTCCCATGTGTATGCAGTCA 2491
    |||||
DB 126 GGCAGCCAAAAAGTTTCAGAAAGGCGGAAAGACATGACTCCCATGTGTATGCAGTCA 185
    |||||
OY 2482 TCGAGGACACCTTGGTATATGGGCATCTGCTACAGATTCCAGCGGCTCCTTCCTGCAGC 2551
    |||||
DB 186 TCGAGGACACCTTGGTATATGGGCATCTGCTACAGATTCCAGCGGCTCCTTCCTGCAGC 245
    |||||
OY 2552 CAGAGGTGGACACTACCGGCGGTTCCAGGGGACCATGGGGGTCTGTCTCCCTCCCCAC 2611
    |||||
DB 246 CAGAGGTGGACACTACCGGCGGTTCCAGGGGACCATGGGGGTCTGTCTCCCTCCCCAC 305
    |||||
OY 2612 CCACCATATGCTCCAGGGGCGCCAACTGCAAAAGTTGGCCA CTGAGAGGCCACCTCTCGCT 2671
    |||||
DB 306 CCACCATATGCTCCAGGGGCGCCAACTGCAAAAGTTGGCCA CTGAGAGGCCACCTCTCGCT 365
    |||||
OY 2672 CCCCTCTGAGTCTGAGAGTGAACCGTACACCTTCTCCCATCCCAACAATGGGGATGTAA 2731
    |||||
DB 366 CCCCTCTGAGTCTGAGAGTGAACCGTACACCTTCTCCCATCCCAACAATGGGGATGTAA 425
    |||||
OY 2732 GCAGCAGGACACAGACATTCCTTACTGAACTCAGGAGGCCCATGGAGCCAGCAGAAAT 2791
    |||||
DB 426 GCAGCAGGACACAGACATTCCTTACTGAACTCAGGAGGCCCATGGAGCCAGCAGAAAT 485
    |||||
OY 2792 AA 2793
    ||
DB 486 AA 487
```

Search completed: February 20, 2004, 04:34:29
Job time : 1489 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2004, 02:41:05 ; Search time 21696 Seconds
(without alignments)
11620.853 Million cell updates/sec

Title: US-09-899-569A-3
Perfect score: 6163
Sequence: 1 ccaacgcgcgaatggggagtc.....agtcgcgcaaaaaaaaaa 6163

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 2045481386 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database :

GenEmbl:
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_by:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pac:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_nub:*
34: em_hcg_pln:*
35: em_hcg_rtd:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_by:*
39: em_hcg_hum:*
40: em_hcg_mus:*
41: em_hcg_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6163	100.0	6163	6 AX353605	AX353605 Sequence
2	5963	96.8	5963	6 AY026461	AY026461 Homo sapi
3	5395	87.5	5897	6 AX353603	AX353603 Sequence
4	4996	81.1	5978	6 AF468010	AF468010 Homo sapi
5	4786	77.7	5573	6 BD160235	BD160235 Primer fo
6	4786	77.7	5573	6 AK023834	AK023834 Homo sapi
7	3268	53.0	163115	9 AC105902	AC105902 Homo sapi
8	3268	53.0	200956	9 AC104165	AC104165 Homo sapi
9	3004	48.7	3224	9 AK026622	AK026622 Homo sapi
10	1703	27.6	2113	9 AK026028	AK026028 Homo sapi
11	1545	25.1	2209	6 AR243778	AR243778 Sequence
12	1177	19.1	1241	9 AK026187	AK026187 Homo sapi
13	1159	18.8	1410	9 BC021099	BC021099 Homo sapi
14	1117	18.1	1376	6 AX464028	AX464028 Sequence
15	909	14.7	976	9 AK026329	AK026329 Homo sapi
16	493	8.0	544	6 AX192795	AX192795 Sequence
17	477	7.7	541	6 AX261684	AX261684 Sequence
18	453	7.4	535	6 AX260742	AX260742 Sequence
19	435	7.1	636	6 AX261015	AX261015 Sequence
20	434	7.0	3301	6 AX576112	AX576112 Sequence
21	404	6.6	659	6 BD150371	BD150371 Primer fo
22	383	6.2	561	6 BD155313	BD155313 Primer fo
23	364	5.9	173977	9 AC010170	AC010170 Homo sapi
24	279	4.5	400	6 AX203285	AX203285 Sequence
25	174	2.8	174	6 AR251796	AR251796 Sequence
26	140	2.3	262	6 AX261252	AX261252 Sequence
27	68	1.1	104531	9 HSDJ303A1	AL096800 Human DNA
28	68	1.1	153241	9 AC023510	AC023510 Homo sapi
29	68	1.1	162589	2 AC015717	AC015717 Homo sapi
30	68	1.1	173585	9 AC090527	AC090527 Homo sapi
31	68	1.1	218724	2 AC025889	AC025889 Homo sapi
32	67	1.1	73775	9 AP003901	AP003901 Homo sapi
33	67	1.1	107469	9 HSB29H4	AL078475 Homo sapi
34	67	1.1	119182	9 HSB62L20	AL050302 Homo sapi
35	67	1.1	162364	9 AP005213	AP005213 Homo sapi
36	67	1.1	190690	2 AL627234	AL627234 Homo sapi
37	67	1.1	340000	9 H821C003	AL163203 Homo sapi
38	64	1.0	155304	9 AC055845	AC055845 Homo sapi
39	64	1.0	174187	2 AP002792	AP002792 Homo sapi
40	64	1.0	215441	9 AP002800	AP002800 Homo sapi
41	64	1.0	215647	2 AC068591	AC068591 Homo sapi
42	63	1.0	724	6 AX417709	AX417709 Sequence
43	63	1.0	106605	9 AL353691	AL353691 Human DNA
44	63	1.0	132290	9 AL627402	AL627402 Human DNA
45	63	1.0	144735	2 AL161458	AL161458 Homo sapi

ALIGNMENTS

RESULT 1
AX353605
LOCUS AX353605
DEFINITION Sequence 3 from Patent WO0204508.
ACCESSION AX353605
VERSION AX353605.1 GI:18618678
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Schwefler,N., Scherl-Mostagier,M., Sommergruber,W. and Abseher,R.
TITLE Tumour-associated antigen (b345), characterised by an amino acid
sequence as in seq. Id. No. 4

Pred. No. is the number of results predicted by chance to have a

[illegible]

OY	541	AATACATTTCGATAGAGATCCAGAAAAAATTTGACATGTAATGTCAGGCCATCATGCTCTTT	600
Db	541	AATACATTTCGATAGAGATCCAGAAAAAATTTGACATGTAATGTCAGGCCATCATGCTCTTT	600
OY	601	GGGAGGTTGACGTTTCAGCCCTGCACATCGTTGTGCTACCCCTCAACAGAACTTTCAATC	660
Db	601	GGGAGGTTGACGTTTCAGCCCTGCACATCGTTGTGCTACCCCTCAACAGAACTTTCAATC	660
OY	661	TGGATGTCMAAGCTCAATAAGAGATGGGTTTAGAGCTGCAGTTTTCATTCCTCCGGCTG	720
Db	661	TGGATGTCMAAGCTCAATAAGAGATGGGTTTAGAGCTGCAGTTTTCATTCCTCCGGCTG	720
OY	721	AGGAGATCCGTCGGGGTGAAGAGCTGGCCAGACGAGATCACTCACTCCATCAGGGCCGA	780
Db	721	AGGAGATCCGTCGGGGTGAAGAGCTGGCCAGACGAGATCACTCACTCCATCAGGGCCGA	780
OY	781	ATCGATCCACCGTGTGTCAGAGATCGAAACCTTTGTCAGCAATGGACATGTGTCCCGATC	840
Db	781	ATCGATCCACCGTGTGTCAGAGATCGAAACCTTTGTCAGCAATGGACATGTGTCCCGATC	840
OY	841	AAGATGCAAGAGAGATGAAAAATGGCCTTTCACCTCCATGATTCACCCAGAAATGTC	900
Db	841	AAGATGCAAGAGAGATGAAAAATGGCCTTTCACCTCCATGATTCACCCAGAAATGTC	900
OY	901	TCCGGCTTTCAGCATTTGCAAAACCGCTCATCTATAAAACGTCTGTGCATCATGATCTGTG	960
Db	901	TCCGGCTTTCAGCATTTGCAAAACCGCTCATCTATAAAACGTCTGTGCATCATGATCTGTG	960
OY	961	TTTGAAGGTGAAGGCTCAGCAACCTCTATATGTCTGCACTAACCCAGAAAGCTTCCCTGAG	1020
Db	961	TTTGAAGGTGAAGGCTCAGCAACCTCTATATGTCTGCACTAACCCAGAAAGCTTCCCTGAG	1020
OY	1021	GATGAGCTCATGAGAGTGGAGTTTGTGTGTCCTGCAACCTGCGGGGCGAGCGTCTCCTTC	1080
Db	1021	GATGAGCTCATGAGAGTGGAGTTTGTGTGTCCTGCAACCTGCGGGGCGAGCGTCTCCTTC	1080
OY	1081	CTCAACTTCAAACCTCTTCCAACTGTGAGAGAAAGGAGCGGGTTGAATCTACATCCCG	1140
Db	1081	CTCAACTTCAAACCTCTTCCAACTGTGAGAGAAAGGAGCGGGTTGAATCTACATCCCG	1140
OY	1141	GGCTTCACACCAACCCCGGAGGTTCMAACTGTGAGGACAAGACGCTGGGAAATATGCGC	1200
Db	1141	GGCTTCACACCAACCCCGGAGGTTCMAACTGTGAGGACAAGACGCTGGGAAATATGCGC	1200
OY	1201	GGGAATTTCAACCTCTCTGTGCAGGCTGTGACCAAGATCCCAAGTCCAGGATCTTC	1260
Db	1201	GGGAATTTCAACCTCTCTGTGCAGGCTGTGACCAAGATCCCAAGTCCAGGATCTTC	1260
OY	1261	CGGCTGCAGTTCCAAAGTTTGTGTCMAACATCCAAATGAAAGCAATTAATCTACGTG	1320
Db	1261	CGGCTGCAGTTCCAAAGTTTGTGTCMAACATCCAAATGAAAGCAATTAATCTACGTG	1320
OY	1321	GTTGACCTTGAGTATGAGCGACGACATGTCATCAACATCGAGGCAAGGCCCGTCAACAG	1380
Db	1321	GTTGACCTTGAGTATGAGCGACGACATGTCATCAACATCGAGGCAAGGCCCGTCAACAG	1380
OY	1381	AGCCGCAAGTTTGTCCCTGGCTGTTTGTGTGTCATGAATCTCGGACCTGCAGTAGCAAC	1440
Db	1381	AGCCGCAAGTTTGTCCCTGGCTGTTTGTGTGTCATGAATCTCGGACCTGCAGTAGCAAC	1440
OY	1441	CTCACCTGCATCTGGCTCCAAACACAAATCTCTTCTTTGTGATCTGACAGCT	1500
Db	1441	CTCACCTGCATCTGGCTCCAAACACAAATCTCTTCTTTGTGATCTGACAGCT	1500
OY	1501	CTGAGATGATATGGAAGAAAAACATTAAGTCGACACACACCGGTCCTGCAAGAAAGAA	1560
Db	1501	CTGAGATGATATGGAAGAAAAACATTAAGTCGACACACACCGGTCCTGCAAGAAAGAA	1560
OY	1561	TCTCACTCACTCCAGGTGCCAAGTCAATCTCCACTGCTGCTGTGAGCTGCATGACTTC	1620
Db	1561	TCTCACTCACTCCAGGTGCCAAGTCAATCTCCACTGCTGCTGTGAGCTGCATGACTTC	1620

QY	1521	CCCTGGAAAGCTGCTGGTGGCCCAAGAGACAGGCTCAGCCGTGGTGGTGGCCAGCCCAAG	1680
Db	1621	TCCTGGAAAGCTGCTGGTGGCCCAAGAGACAGGCTCAGCCGTGGTGGTGGCCAGCCCAAG	1680
QY	1681	CTGCAGACGACATACACAGAGAACCCCTGCACACACAGCTTCAGCTACCTTGTCGACGT	1740
Db	1681	CTGCAGACGACATACACAGAGAACCCCTGCACACACAGCTTCAGCTACCTTGTCGACGT	1740
QY	1741	GCCATACCCAGACGACGACTGTACTTGCGCTCTTCTGCCCCGGAGGCTCTATCAACAG	1800
Db	1741	GCCATACCCAGACGACGACTGTACTTGCGCTCTTCTGCCCCGGAGGCTCTATCAACAG	1800
QY	1801	ATCCAGGTGAAGCAGAAACATCTCGGTACCCCTTGCGACCTTTGGCCCCCAGCTTCCACAA	1860
Db	1801	ATCCAGGTGAAGCAGAAACATCTCGGTACCCCTTGCGACCTTTGGCCCCCAGCTTCCACAA	1860
QY	1861	GAGCCCTTCAGAGCAGGCTGTGACGGTGTCCCTTATACCTTATTTCAAGAGAGAGCGTT	1920
Db	1861	GAGCCCTTCAGAGCAGGCTGTGACGGTGTCCCTTATACCTTATTTCAAGAGAGAGCGTT	1920
QY	1921	TTTCACGGGTGACCCCTGACACAAAAAGCAGGTCTACCTGAGACCCCCAACTGGGACCG	1980
Db	1921	TTTCACGGGTGACCCCTGACACAAAAAGCAGGTCTACCTGAGACCCCCAACTGGGACCG	1980
QY	1981	GGCCTGCGCATCCCTCACTCTGTGTCTCTGGAACATCAGCGTGGCCCAAGACACAGTGGCC	2040
Db	1981	GGCCTGCGCATCCCTCACTCTGTGTCTCTGGAACATCAGCGTGGCCCAAGACACAGTGGCC	2040
QY	2041	TGCGCTGACCTTCTTTAAGAGACCGGAGCGGCGCTGGCTGCGCAGACAGGAGCGCGCATTCATG	2100
Db	2041	TGCGCTGACCTTCTTTAAGAGACCGGAGCGGCGCTGGCTGCGCAGACAGGAGCGCGCATTCATG	2100
QY	2101	ATCATCCAGAGACGACCGGACCCGGGCTGAGAGATCTTCAAGCTGAGACGAGATGTGCTC	2160
Db	2101	ATCATCCAGAGACGACCGGACCCGGGCTGAGAGATCTTCAAGCTGAGACGAGATGTGCTC	2160
QY	2161	CCCAAAGCCAGCTTCCACATCAACAGCTTCTGGGTCAAACATCTCTAACTGACGCCCAAG	2220
Db	2161	CCCAAAGCCAGCTTCCACATCAACAGCTTCTGGGTCAAACATCTCTAACTGACGCCCAAG	2220
QY	2221	AGGGGCAAGCAGCTGAGACCTGCTCTTCTCGGGTGACCTTACCCCAAGACCTGTGACCTTG	2280
Db	2221	AGGGGCAAGCAGCTGAGACCTGCTCTTCTCGGGTGACCTTACCCCAAGACCTGTGACCTTG	2280
QY	2281	ACTGTCACTCTCATCTGACGAGCGGTGGAGGTGAGATCTTACTGTCTGTCTGCGCTCGGCTC	2340
Db	2281	ACTGTCACTCTCATCTGACGAGCGGTGGAGGTGAGATCTTACTGTCTGTCTGCGCTCGGCTC	2340
QY	2341	ATCATTTGCTGTGTGAAAAAGAGAAAAAGAAACAAACAGGGCCCCGCTGTGGGTATC	2400
Db	2341	ATCATTTGCTGTGTGAAAAAGAGAAAAAGAAACAAACAGGGCCCCGCTGTGGGTATC	2400
QY	2401	TACAAATGGCAACATCAATATCTGAGATGCCGAGCGACCCAAAAAGTTTCAGAAAGGGCGA	2460
Db	2401	TACAAATGGCAACATCAATATCTGAGATGCCGAGCGACCCAAAAAGTTTCAGAAAGGGCGA	2460
QY	2461	AAGAGCAATGACTCCCATGTGTATGCAAGTCAATCAGAGACACATGTGATATGAGCATCTG	2520
Db	2461	AAGAGCAATGACTCCCATGTGTATGCAAGTCAATCAGAGACACATGTGATATGAGCATCTG	2520
QY	2521	CTACAGAGATTCACAGCGGCTCCTTCCTGACGCAAGAGTGAACACTACCGGCGTTCCAG	2580
Db	2521	CTACAGAGATTCACAGCGGCTCCTTCCTGACGCAAGAGTGAACACTACCGGCGTTCCAG	2580
QY	2581	GGCACCATGGGGGTCTGTCTCTCTCCCAACCAACATATGCTTCAAGGGCCCCAACTGCA	2640
Db	2581	GGCACCATGGGGGTCTGTCTCTCTCCCAACCAACATATGCTTCAAGGGCCCCAACTGCA	2640
QY	2641	AAGTTGGGCACTGAGAGACCACTCTGCTGCCCTCCTGTAGTCTGAGAGTGAACCGTAC	2700
Db	2641	AAGTTGGGCACTGAGAGACCACTCTGCTGCCCTCCTGTAGTCTGAGAGTGAACCGTAC	2700
QY	2701	ACCTTCTCCATCCCAACATGGGGATGTAGAAGCAAGACACAGACATTTCCCTTACTG	2760

Db	2701	ACCTTCTCCATCCCAACAAATGGGGATGTAAAGCAGACAGACACAGCATTCCTCTTAAGT	2760
Qy	2761	AACACTGAGAGCCCATGAGCCGACGAGAAATAACTTGATCATTCACGACGCTTGTGCTA	2820
Db	2761	AACACTGAGAGCCCATGAGCCGACGAGAAATAACTTGATCATTCACGACGCTTGTGCTA	2820
Qy	2821	GTTTCATTAAGCAGAGGCACCTGAGACACCCGTCGGTTCCTAACCAAAATTCCTAAAGAA	2880
Db	2821	GTTTCATTAAGCAGAGGCACCTGAGACACCCGTCGGTTCCTAACCAAAATTCCTAAAGAA	2880
Qy	2881	GAGGAATTAATCAGAAAGAAACAGAGAGAGGTTTTCTGGAACCGGCCAATTCCATATTCG	2940
Db	2881	GAGGAATTAATCAGAAAGAAACAGAGAGAGGTTTTCTGGAACCGGCCAATTCCATATTCG	2940
Qy	2941	TCAGTGGACCTCATTTCTAAGGGCAGACATTTGAAATGATGAATTCCAATCTGATACAGT	3000
Db	2941	TCAGTGGACCTCATTTCTAAGGGCAGACATTTGAAATGATGAATTCCAATCTGATACAGT	3000
Qy	3001	CATACACAGCTCATGTGCTCTCAACTTAAGCTGTGCGGTTTAGCCAGGCTGTATGAGAGG	3060
Db	3001	CATACACAGCTCATGTGCTCTCAACTTAAGGCTGTGCGGTTTAGCCAGGCTGTATGAGAGG	3060
Qy	3061	AGAGAGGCTGAGTCACTTAGCAATAGGGTTGACGAAAGCCCTGGATTCAAGATGTTAAAC	3120
Db	3061	AGAGAGGCTGAGTCACTTAGCAATAGGGTTGACGAAAGCCCTGGATTCAAGATGTTAAAC	3120
Qy	3121	AGAGGCTTGGCCCTCTTTCAGAGCAACAAGTTCCAATTCGAAAGAGCTACGAGGTCCCTA	3180
Db	3121	AGAGGCTTGGCCCTCTTTCAGAGCAACAAGTTCCAATTCGAAAGAGCTACGAGGTCCCTA	3180
Qy	3181	CTCTCACTGAGGGTCCCGAGATGMAAACGACATGTGCTCTTTTATTAATTAATTTATTTGG	3240
Db	3181	CTCTCACTGAGGGTCCCGAGATGMAAACGACATGTGCTCTTTTATTAATTAATTTATTTGG	3240
Qy	3241	TGCTTCCTGTGTATTTAAAGATTCGAAATGATTAACAACCTAGCTCTTTTCACTGACCTTA	3300
Db	3241	TGCTTCCTGTGTATTTAAAGATTCGAAATGATTAACAACCTAGCTCTTTTCACTGACCTTA	3300
Qy	3301	GTAATTAACCTCACTAATCTGGTTTGATGCTGGGTTGTGACTTCTAATCAGCCGCTAGAT	3360
Db	3301	GTAATTAACCTCACTAATCTGGTTTGATGCTGGGTTGTGACTTCTAATCAGCCGCTAGAT	3360
Qy	3361	AAACGTGTGCTGTCGCCCAAGGTGTGGGAATTAATTTACAATCTGTCCAAACAGAAAGA	3420
Db	3361	AAACGTGTGCTGTCGCCCAAGGTGTGGGAATTAATTTACAATCTGTCCAAACAGAAAGA	3420
Qy	3421	ATGTGTGTGTTTGAAGCATTTGACACATATCTGCTTTGATTAAGAGCTTCCTGATTTCTC	3480
Db	3421	ATGTGTGTGTTTGAAGCATTTGACACATATCTGCTTTGATTAAGAGCTTCCTGATTTCTC	3480
Qy	3481	TAGGTGCGGTTGCTGTTATCCCATTTGGAAATTCATCTTGAATCCATGTCTCTAATAGT	3540
Db	3481	TAGGTGCGGTTGCTGTTATCCCATTTGGAAATTCATCTTGAATCCATGTCTCTAATAGT	3540
Qy	3541	CCTAGCAATAGAGAAATTTCTCAAGTTTCCATGTGCGGTTCTCTTAGCTGACGACATA	3600
Db	3541	CCTAGCAATAGAGAAATTTCTCAAGTTTCCATGTGCGGTTCTCTTAGCTGACGACATA	3600
Qy	3601	CTTTGACATTTAAAGAGAAATTTAGAGAAATTTCTCATCTCTTAATAATGTTTAAATATA	3660
Db	3601	CTTTGACATTTAAAGAGAAATTTAGAGAAATTTCTCATCTCTTAATAATGTTTAAATATA	3660
Qy	3661	TACCAAAACAGTGGCCCTGCTGATTAAGTTTCTGTGGCACTGGCAACCATTAATCTGGTAG	3720
Db	3661	TACCAAAACAGTGGCCCTGCTGATTAAGTTTCTGTGGCACTGGCAACCATTAATCTGGTAG	3720
Qy	3721	CTTAATAAAACAACATTAAGCTTAATAGTCTGTGGGATCAGAAATTCAAAATGGAATGTCCCT	3780
Db	3721	CTTAATAAAACAACATTAAGCTTAATAGTCTGTGGGATCAGAAATTCAAAATGGAATGTCCCT	3780
Qy	3781	GAATGAATAATCAGGTGTCAAGCAGAGCTGTGCTCTTTCTGAAGGCTTTAGGAGAGAGCCG	3840
Db	3781	GAATGAATAATCAGGTGTCAAGCAGAGCTGTGCTCTTTCTGAAGGCTTTAGGAGAGAGCCG	3840

Dh 3781 GAATGAAATCAAGGTCTAGACAGAGCTGTGCTCTTTCTGAAAGGCTTGAAGAAAGCCG 3840
Qy 3841 GTTCCTTGCCATTTCAAGCTTCTAGAGGCTGCTGCAATTCAGAGGCTTCAGTGGCTGTC 3900
Dh 3841 GTTCCTTGCCATTTCAAGGCTTCTAGAGGCTGCTGCAATTCAGAGGCTTCAGTGGCTGTC 3900
Qy 3901 AAGCTTTTCTCAGATGGCATCACTGTGACACTGAGCCCTCCACTTCCCTTTTGAATTAC 3960
Dh 3901 AAGCTTTTCTCAGATGGCATCACTGTGACACTGAGCCCTCCACTTCCCTTTTGAATTAC 3960
Qy 3961 AAAGCCACAGGAAGATCCAGGATATCTCTCATCTAAAGATCTTCATCATCTCTGA 4020
Dh 3961 AAAGCCACAGGAAGATCCAGGATATCTCTCATCTAAAGATCTTCATCATCTCTGA 4020
Qy 4021 AGAGCCTTTTGCATCAAGACAAATAGCAAGGTGGGAGATTAGACAGAGCACTT 4080
Dh 4021 AGAGCCTTTTGCATCAAGACAAATAGCAAGGTGGGAGATTAGACAGAGCACTT 4080
Qy 4081 TGGGGTGTGTTATTTCTGCTTACCACTCTTCTGCACTGACTCCCAAGAGAGCTA 4140
Dh 4081 TGGGGTGTGTTATTTCTGCTTACCACTCTTCTGCACTGACTCCCAAGAGAGCTA 4140
Qy 4141 CAAATGATCTGGCGACAGGAGTGTGTTAGCTGGAGCTTAACTTAATAAAA 4200
Dh 4141 CAAATGATCTGGCGACAGGAGTGTGTTAGCTGGAGCTTAACTTAATAAAA 4200
Qy 4201 ACCCCAGATCAGAAAGATCTGGCCATCTGGGCTCAATTCATCACTTACCAACATGAGC 4260
Dh 4201 ACCCCAGATCAGAAAGATCTGGCCATCTGGGCTCAATTCATCACTTACCAACATGAGC 4260
Qy 4261 TGGAGCTGGGACACAGCTCTGCTTTTGAAGAGGGGTCCACTTACACAGTCAACACAGC 4320
Dh 4261 TGGAGCTGGGACACAGCTCTGCTTTTGAAGAGGGGTCCACTTACACAGTCAACACAGC 4320
Qy 4321 CCACATAGGCCCTTACCTTCCACAAATGAGGCTAAGTGTGTTTCTTACTAGATCAATG 4380
Dh 4321 CCACATAGGCCCTTACCTTCCACAAATGAGGCTAAGTGTGTTTCTTACTAGATCAATG 4380
Qy 4381 CCCCTCAGGTTGATTTATTTGTAATGAAAGAAAGAGCTGGATTAATCTTAATCAG 4440
Dh 4381 CCCCTCAGGTTGATTTATTTGTAATGAAAGAAAGAGCTGGATTAATCTTAATCAG 4440
Qy 4441 TGAATGACCATGAGACCAATGTGTCTCACTTACCCTTTTCTTTTCTTTTCT 4500
Dh 4441 TGAATGACCATGAGACCAATGTGTCTCACTTACCCTTTTCTTTTCTTTTCT 4500
Qy 4501 TTTTCTTTTCTTTTAAATGTAAGAGAGATCTCATTTGTTGCTTGAAGCTGAGTGCAG 4560
Dh 4501 TTTTCTTTTCTTTTAAATGTAAGAGAGATCTCATTTGTTGCTTGAAGCTGAGTGCAG 4560
Qy 4561 TGGCGCAATCTCGGCTCACTGCAACCTTCTGCTCTGAGGCTCAAGCAATTTCTCCACTC 4620
Dh 4561 TGGCGCAATCTCGGCTCACTGCAACCTTCTGCTCTGAGGCTCAAGCAATTTCTCCACTC 4620
Qy 4621 AGCTTCCCAATAGCTGGATCACTGGCAAAACCAACATGCCCACATATTTTGTATTT 4680
Dh 4621 AGCTTCCCAATAGCTGGATCACTGGCAAAACCAACATGCCCACATATTTTGTATTT 4680
Qy 4681 TTTTGTAGACAGGGTTTCAACCATGTTGCCAGGCTGTCTCAACCTCCCTGGGCTCAAGC 4740
Dh 4681 TTTTGTAGACAGGGTTTCAACCATGTTGCCAGGCTGTCTCAACCTCCCTGGGCTCAAGC 4740
Qy 4741 AATCTCTGCTCGGCTCCCAAAAGTGTGGAGTTTACAGATGAGCCACCGCATCCAG 4800
Dh 4741 AATCTCTGCTCGGCTCCCAAAAGTGTGGAGTTTACAGATGAGCCACCGCATCCAG 4800
Qy 4801 CCCCAACCTCTATTATACCAATTAACCTGCCCCAGTAACGTGAGCTTTTGTCTCAC 4860
Dh 4801 CCCCAACCTCTATTATACCAATTAACCTGCCCCAGTAACGTGAGCTTTTGTCTCAC 4860
Qy 4861 CCTGTCTGATCTGGAAGAGAGGATTAATGTTATAGCTTGTACACAGTCCCAAGTT 4920
Dh 4861 CCTGTCTGATCTGGAAGAGAGGATTAATGTTATAGCTTGTACACAGTCCCAAGTT 4920

Qy 4921 CAATATTTCTGGGCAAAAACCTTCTTCAAAAATTAATGTAATTTGATTAATGA 4980
Dh 4921 CAATATTTCTGGGCAAAAACCTTCTTCAAAAATTAATGTAATTTGATTAATGA 4980
Qy 4981 AATCACTTGGAAATGCAACCGCTCACTTGTGCAATGGCAATAATGAAGAAATTTTA 5040
Dh 4981 AATCACTTGGAAATGCAACCGCTCACTTGTGCAATGGCAATAATGAAGAAATTTTA 5040
Qy 5041 TAGTCTCCAAATGAGCGTGTACTGCAAGACCTTTGAACACTTTCCAGAGATAGATAT 5100
Dh 5041 TAGTCTCCAAATGAGCGTGTACTGCAAGACCTTTGAACACTTTCCAGAGATAGATAT 5100
Qy 5101 TTAAGTCAATGCTTGGCGTGTGCTATGAGCACTTTCCCTTGTAAAGTCTGCTTGC 5160
Dh 5101 TTAAGTCAATGCTTGGCGTGTGCTATGAGCACTTTCCCTTGTAAAGTCTGCTTGC 5160
Qy 5161 CCAATGACCCCTTGGCTTGTAGAGCCGAGATGTGACCTTGCATAAAGGCGCAAGAGGG 5220
Dh 5161 CCAATGACCCCTTGGCTTGTAGAGCCGAGATGTGACCTTGCATAAAGGCGCAAGAGGG 5220
Qy 5221 CTGGGCTTCTTCCCTCACTGAAAGCCCTTAATTTGAATTCATGCTGTGAGCCCTAGC 5280
Dh 5221 CTGGGCTTCTTCCCTCACTGAAAGCCCTTAATTTGAATTCATGCTGTGAGCCCTAGC 5280
Qy 5281 CCTCAATCTGGAATTTCCCAACCTCCAGCCCTTCCAGACAGACTAGGTGCCCTGC 5340
Dh 5281 CCTCAATCTGGAATTTCCCAACCTCCAGCCCTTCCAGACAGACTAGGTGCCCTGC 5340
Qy 5341 AATCAACCAAGGTGGAGATGGGCTTCTTGGGCTGAGTCTATGACCATCAACGACAT 5400
Dh 5341 AATCAACCAAGGTGGAGATGGGCTTCTTGGGCTGAGTCTATGACCATCAACGACAT 5400
Qy 5401 CACTGTGCTGTGCAAGACACCAAGTGGCAATTTCTTCAACTGAGGCTCAAAACTCC 5460
Dh 5401 CACTGTGCTGTGCAAGACACCAAGTGGCAATTTCTTCAACTGAGGCTCAAAACTCC 5460
Qy 5461 TGAACAAGTGTGCTCTGAGACCAAGATTTCTGAGCTGTGCTTCACTGAAGGGGC 5520
Dh 5461 TGAACAAGTGTGCTCTGAGACCAAGATTTCTGAGCTGTGCTTCACTGAAGGGGC 5520
Qy 5521 CCAAGCTGAGGAACCTGCTCTTTCTTAAAGCCAGGCCCACTTACATTAACAT 5580
Dh 5521 CCAAGCTGAGGAACCTGCTCTTTCTTAAAGCCAGGCCCACTTACATTAACAT 5580
Qy 5581 TCAGGGTCACTGGAACAAGTGAAGTGCATTTGTTGAAGCTTACATGACGCCACT 5640
Dh 5581 TCAGGGTCACTGGAACAAGTGAAGTGCATTTGTTGAAGCTTACATGACGCCACT 5640
Qy 5641 GCTCATCAAGTGTGTGCTGATGCTTACGAGAGGCGCAGCGCATGCAAGACTGTCT 5700
Dh 5641 GCTCATCAAGTGTGTGCTGATGCTTACGAGAGGCGCAGCGCATGCAAGACTGTCT 5700
Qy 5701 AATGCTGTGTATTTGCAACAAGGAAAGGCTTCAAGAAAGTCACTGGAGCAAGA 5760
Dh 5701 AATGCTGTGTATTTGCAACAAGGAAAGGCTTCAAGAAAGTCACTGGAGCAAGA 5760
Qy 5761 CAAGCCACCGGACATGAGCTTGTGAAGGTTAGAGAGCTGTGTGTGATCTGACAT 5820
Dh 5761 CAAGCCACCGGACATGAGCTTGTGAAGGTTAGAGAGCTGTGTGTGATCTGACAT 5820
Qy 5821 GCTTCACTGGAATTAATTTATTTATTTGAGATCTTTTAAAGTGCATTTTATTCAT 5880
Dh 5821 GCTTCACTGGAATTAATTTATTTATTTGAGATCTTTTAAAGTGCATTTTATTCAT 5880
Qy 5881 CTGTGCTTTAATTAACAAGATACCAAAAAACAGATCAAGCTGTTTAAGTCTTGG 5940
Dh 5881 CTGTGCTTTAATTAACAAGATACCAAAAAACAGATCAAGCTGTTTAAGTCTTGG 5940
Qy 5941 CTAATTTGCCCTGTGTTAGTGAAGCCCGGTTTCCAGTTTGTGACTGTGACAGGCTC 6000
Dh 5941 CTAATTTGCCCTGTGTTAGTGAAGCCCGGTTTCCAGTTTGTGACTGTGACAGGCTC 6000

QY 6001 ASCATGGCTCAGACAGATCTGTCTTAATTGTGATGATACAGAAAGCCAGGCTTTGGG 6060
DB 6001 ACCATGGGCTCAGACAGATCTGTCTTAATTGTGATGATACAGAAAGCCAGGCTTTGGG 6060
QY 6061 AATCAAGTCTTCTCTTCATTTGATGCGCGTGCATGTGTGAAGAGATGTTTTTGTCC 6120
DB 6061 AATCAAGTCTTCTCTTCATTTGATGCGCGTGCATGTGTGAAGAGATGTTTTTGTCC 6120
QY 6121 GGAATATATATATATAGTCTTGAGTCTGCGCCAAAAA 6163
DB 6121 GGAATATATATATATAGTCTTGAGTCTGCGCCAAAAA 6163

RESULT 2
AY026461 5963 bp mRNA linear PRI 17-AUG-2001
LOCUS Homo sapiens CUB domain containing protein 1 (CDCP1) mRNA, complete
DEFINITION
ACCESSION AY026461
VERSION AY026461.1 GI:14328878
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 5963)
Scherl-Mostagguer, W., Sommergruber, W., Abseher, R., Hauptmann, R.,
Ambros, P. and Schweizer, N.
Identification of a novel gene, CDCP1, overexpressed in human
colorectal cancer
Oncogene 20 (32), 4402-4408 (2001)
JOURNAL MEDLINE 21359860
PUBMED 11466621
TITLE 2 (bases 1 to 5963)
Scherl-Mostagguer, W., Sommergruber, W., Abseher, R., Hauptmann, R.,
Ambros, P. F. and Schweizer, N.
Direct Submision
SUBMITTED (31-JAN-2001) Exploratory Research, Boehringer Ingelheim,
Austria, Dr. Boehringergasse 5-11, Vienna 1121, Austria
FEATURES
source
1..5963
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="3"
/map="3p23-p21"
/cell_line="A549; Calu 6"
1..5963
/gene="CDCP1"
1..82
/gene="CDCP1"
83..2593
/gene="CDCP1"
/note="transmembrane protein"
/codon_start=1
/product="CUB domain containing protein 1"
/protein_id="AAK02058.1"
/db_xref="GI:14328878"
/translation="MAGLNGVSIALLGVLLGAARLPRGAAFEIALPRESNITVLI
KGTPTLAKPYIVISKRIIMLSIKSGERIVFTPSQSPENHPIEIOKXIDCMG
PPGPEVQLQSTSLPTLNFIMDKVSKSIGLELOSIPLRLOIGSGESGQVY
HSISGRIDATVVRIGTFCNGTVSRIMQEGVKMALHLMFHRNVRSGSIANRSIK
RJCITESVPEGSGATIMSNYBEPGPEELMTQFVVAHRAVSPLNPLNSCER
KEERBYIIPGSTTPPEVFKEDKQGNMAGFNLSLQCDDDAGSPGILRLQFOVLV
QHPQENSKIYVVDLSNERAMSLTEPRVVKOSRKFVPCFVCLSRSSNLTLTG
SHKISFLCDDLTRLMNVKETSCTDHRVCQKSYSLQSPDILHLPELHDFMKG
LVPKRLSLVLPAAQLQHTHEKPCNTSFSLVASAIPSOQLYSGSPGSGIKQIQ
VKNQISVTLRTFAPSFQOASRQGLVSRIPYKEGAVTTPDRTSKYLYLTPMMDR
GLPSLTVSNWNI SVPRDOVACLTPKEKSGVCCGGRAMIIQDQRTAEELPSLDED
VLPRGSHHSHFVNINSCSPISGKOLDLFSYTLTPRVLDLTVILIAVGGVLLS
ALGLICGVKKKKKTKNKPAGVINGNINTEMPROPKPFQGRKNDKSHVAIVEDT
NVYGHLDSSGSLFQPEVDYIARPFQGTGVCPPSPPTCSAPAKPALKATEBPPPSP
PESEBPTFTFHPNNGVSSKTDIPLANTQEMPAE"

3'UTR 2594..5963
BASE COUNT 1463 a 1634 c 1392 g 1474 t
ORIGIN
Query Match 96.8%; Score 5963; DB 9; Length 5963;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5963; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 201 GCGACGTGGCGGGCTTGGGGGCGCTGGGCGCGGCTCCGACGCTGTTCCCAAC 260
DB 1 GCGACGTGGCGGGCTTGGGGGCGCTGGGCGCGGCTCCGACGCTGTTCCCAAC 60
QY 261 GAGGCGAGGCGCTCCCGAGTCATGACCGGCTGAACTGCGGGCTCTATGACATGCT 320
DB 61 GAGGCGAGGCGCTCCCGAGTCATGACCGGCTGAACTGCGGGCTCTATGACATGCT 120
QY 321 AGGGGTTCTGCTGCTGGGGTGGCGCGGCTGCGCGCGGCGGAGCTTTGAGATTGC 380
DB 121 AGGGGTTCTGCTGCTGGGGTGGCGCGGCTGCGCGCGGCGGAGCTTTGAGATTGC 180
QY 381 TCTGCGACGAGAAAGCAATTACAGTTCTCAATAAGCTGGGGACCCGACCTGCTGGC 440
DB 181 TCTGCGACGAGAAAGCAATTACAGTTCTCAATAAGCTGGGGACCCGACCTGCTGGC 240
QY 441 AAAACCTGTTACATGTCATTTCTAAAGACATATAACATGTTGTCATCAAGTCTGG 500
DB 241 AAAACCTGTTACATGTCATTTCTAAAGACATATAACATGTTGTCATCAAGTCTGG 300
QY 501 AGAAAGAAATAGTCTTTACCTTTAGCTGCGCAGATCTTGAATCACTTTGATAGAT 560
DB 301 AGAAAGAAATAGTCTTTACCTTTAGCTGCGCAGATCTTGAATCACTTTGATAGAT 360
QY 561 CCAGAAATATATGACTGATGTCAGGCCAGTCTTTGGGGAGTTACGCTTCAGGC 620
DB 361 CCAGAAATATATGACTGATGTCAGGCCAGTCTTTGGGGAGTTACGCTTCAGGC 420
QY 621 CTGCACATGCTGTTGGCTCAACCTCAACAGAACTTTCACTTGGAGTCAAACTCATTA 680
DB 421 CTGCACATGCTGTTGGCTCAACCTCAACAGAACTTTCACTTGGAGTCAAACTCATTA 480
QY 681 GAGCATCGGTTTGAAGCTGCACTGTTTCATGCTGCTGAGGAGATCGTCCGGGTGA 740
DB 481 GAGCATCGGTTTGAAGCTGCACTGTTTCATGCTGCTGAGGAGATCGTCCGGGTGA 540
QY 741 GAGCTGCCCAAGCGAGTCACTCATCTCACTCAAGCGGCGGAATGCACTGCTGTGAG 800
DB 541 GAGCTGCCCAAGCGAGTCACTCATCTCACTCAAGCGGCGGAATGCACTGCTGTGAG 600
QY 801 GATCGAAACCTTCTGACAGATGACATGCTGCTCCCGGATCAAGATGCAAGAGAGTGA 860
DB 601 GATCGAAACCTTCTGACAGATGACATGCTGCTCCCGGATCAAGATGCAAGAGAGTGA 660
QY 861 AATGAGCTTACACTCTCCATGAGTTCCACCCGAGAAATGCTTCGCGCTTCAGATTGCA 920
DB 661 AATGAGCTTACACTCTCCATGAGTTCCACCCGAGAAATGCTTCGCGCTTCAGATTGCA 720
QY 921 CCGCTCATCTATAAAGCTGTGTGATCATGCAAGTGTGTTGAGGGTGAAGGCTCAGC 980
DB 721 CCGCTCATCTATAAAGCTGTGTGATCATGCAAGTGTGTTGAGGGTGAAGGCTCAGC 780
QY 981 AACCTGATGCTGCACTACCAAGAGGCTTCCTGAGATGAGTGCATGACAGTGA 1040
DB 781 AACCTGATGCTGCACTACCAAGAGGCTTCCTGAGATGAGTGCATGACAGTGA 840
QY 1041 GTTGTGCTTCTGTCACACTGCGGCGCAGCGCTCTCTCTCAACTTCAACTCTGCA 1100
DB 841 GTTGTGCTTCTGTCACACTGCGGCGCAGCGCTCTCTCTCAACTTCAACTCTGCA 900
QY 1101 CTGTGAGGAGAAAGAGAGCGGGTGAATCTAATCTCCGGGCTCAACCAACCCGA 1160
DB 901 CTGTGAGGAGAAAGAGAGCGGGTGAATCTAATCTCCGGGCTCAACCAACCCGA 960

QY	1161	GGTGTTCAGGCTGAGAGACAAAGCAGCCTTGGGAACATGGCCGGGAAATTCACCTCTCTCT	1220
Db	961	GGGTGTTCAGGCTGAGAGACAAAGCAGCCTTGGGAAACATGGCCGGGAAATTCACCTCTCTCT	1020
QY	1221	GCAAGGCTGTGACCAAGATGCGCCAAAGTGCACAGGGATCCTCCGGCTGCAGTTCCAAAGTTT	1280
Db	1021	GCAAGGCTGTGACCAAGATGCGCCAAAGTGCACAGGGATCCTCCGGCTGCAGTTCCAAAGTTT	1080
QY	1281	GGTCCAAATCACAACAAATGAAAGCAATAAATCTACGTGGTTGACTTGAATATGAGCG	1340
Db	1081	GGTCCAAATCACAACAAATGAAAGCAATAAATCTACGTGGTTGACTTGAATATGAGCG	1140
QY	1341	AGCCATGTCTACTACCATCTGAGGCCAGGCCCCGTCAACACAGCCGCAAGTTTGTCCCTGG	1400
Db	1141	AGCCATGTCTACTACCATCTGAGGCCAGGCCCCGTCAACACAGCCGCAAGTTTGTCCCTGG	1200
QY	1401	CTGTGTTGGTGTCTAGAAATCTCGGAGCTGCAAGTAGAGCAACTCACCCGTACATCTGGCTC	1460
Db	1201	CTGTGTTGGTGTCTAGAAATCTCGGAGCTGCAAGTAGAGCAACTCACCCGTACATCTGGCTC	1260
QY	1461	CAAAACACAAAATCTCTCTTCTCTTTGTGATGATCTGACACGTCTGTGATGATGTGAAAA	1520
Db	1261	CAAAACACAAAATCTCTCTTCTCTTTGTGATGATCTGACACGTCTGTGATGATGTGAAAA	1320
QY	1521	AAACCTAAGCTGTGCACAGCACACCCGGTACTGGCCAAAGAAATCTCACTCCACGGTGGC	1580
Db	1321	AAACCTAAGCTGTGCACAGCACACCCGGTACTGGCCAAAGAAATCTCACTCCACGGTGGC	1380
QY	1581	CAGTACATCCTCCACCTGCGCTGTGGAGCTGACATGACATCTTCTCTGAAAGCTGTGGTGGC	1640
Db	1381	CAGTACATCCTCCACCTGCGCTGTGGAGCTGACATGACATCTTCTCTGAAAGCTGTGGTGGC	1440
QY	1641	CAAGGACAGGCTCAGCCTGTGTGTGTGTGTCACAGCCAGAGAGCTGACAGCATACACAGA	1700
Db	1441	CAAGGACAGGCTCAGCCTGTGTGTGTGTGTCACAGCCAGAGAGCTGACAGCATACACAGA	1500
QY	1701	GAAAGCTGTGCACACACAGCTTCACTTGTGTGGCCAGTGCATATCCAGCCACAGACT	1760
Db	1501	GAAAGCTGTGCACACACAGCTTCACTTGTGTGGCCAGTGCATATCCAGCCACAGACT	1560
QY	1761	GTACTTGGGCTCTCTTCTGCGCCGGAGGCTCTATCAAGACATCCAGGTGAAGCGAAACAT	1820
Db	1561	GTACTTGGGCTCTCTTCTGCGCCGGAGGCTCTATCAAGACATCCAGGTGAAGCGAAACAT	1620
QY	1821	CTCGGTGACCTTGCACACTTTTGCCCCAGCTTCAACAAAGAGGCTTCAGAGCAGGCT	1880
Db	1621	CTCGGTGACCTTGCACACTTTTGCCCCAGCTTCAACAAAGAGGCTTCAGAGCAGGCT	1680
QY	1881	GACGGTGTCTTTATATACCTTATTTCAAAAGAGAGGCGTTTTCAACGATGACCCCTGCACAC	1940
Db	1681	GACGGTGTCTTTATATACCTTATTTCAAAAGAGAGGCGTTTTCAACGATGACCCCTGCACAC	1740
QY	1941	AAAAAGCAAGGTCTACCTTGAGAGACCCCCCACTGGGACCGGGGCTGTGCATCCCTCACCTC	2000
Db	1741	AAAAAGCAAGGTCTACCTTGAGAGACCCCCCACTGGGACCGGGGCTGTGCATCCCTCACCTC	1800
QY	2001	TGTGTCTCTGAAACATCAGCGTGTGCCAGAGACCAAGGTGCTGTGCTGACTTCTTTAAGA	2060
Db	1801	TGTGTCTCTGAAACATCAGCGTGTGCCAGAGACCAAGGTGCTGTGCTGACTTCTTTAAGA	1860
QY	2061	GCGAGCGGCGTGTGTCTGACAGACAGGGCGGCAATTCATGATCATCCAGAGACAGGAC	2120
Db	1861	GCGAGCGGCGTGTGTCTGACAGACAGGGCGGCAATTCATGATCATCCAGAGACAGGAC	1920
QY	2121	CCGGGCTGAGAGATCTTCAGCTTGGACGAGAGATGTGCTCCCCAAAGCCAAAGCTTCACCA	2180
Db	1921	CCGGGCTGAGAGATCTTCAGCTTGGACGAGAGATGTGCTCCCCAAAGCCAAAGCTTCACCA	1980
QY	2181	TCACAGCTTCTGGGGCAACATCTCAACGAGAGCCCAAGAGGCGCAAGAGCAAGCACTAGACCT	2240
Db	1981	TCACAGCTTCTGGGGCAACATCTCAACGAGAGCCCAAGAGGCGCAAGAGCAAGCACTAGACCT	2040
QY	2241	GCTCTTCTCGGTGACACTTATCCCAAGACCTGTGACTTGTGACTGTCACTCTCATCTGCAGC	2300

Db	2041	GCTCTTCCTGGTGCACTTACCCCAAGACCTGGAGCTTGACTGTCACTCTCATCGCAGC	2100
QY	2301	GGTGGAGAGTGAATCTTATGCTGTCTGGCCCTTCGGGCTCATCTTTGCTGTGAAAA	2360
Db	2101	GGTGGAGAGTGAATCTTATGCTGTCTGGCCCTTCGGGCTCATCTTTGCTGTGAAAA	2160
QY	2361	GAAGAAAAAGAGACAACAAGGGCCCCGCTGTGGGTATCTACATGGCAATCAATAC	2420
Db	2161	GAAGAAAAAGAGACAACAAGGGCCCCGCTGTGGGTATCTACATGGCAATCAATAC	2220
QY	2421	TGAGATGCCGAGGAGCGCAAAAAAGTTTCAGAAAGGGCGAAAGGCAATGCTCCATGT	2480
Db	2221	TGAATATGCCGAGGAGCGCAAAAAAGTTTCAGAAAGGGCGAAAGGCAATGCTCCATGT	2280
QY	2481	GTATGCAGTATCGAGAGACACATGTATATAGGCACTGTCTACAGATTTCCAGGGCTC	2540
Db	2281	GTATGCAGTATCGAGAGACACATGTATATAGGCACTGTCTACAGATTTCCAGGGCTC	2340
QY	2541	CTTCTCTGAGCGAGAGGTGACACTTACCGGCGCTTCAGGGCACATGGGGGTCTGTCC	2600
Db	2341	CTTCTCTGAGCGAGAGGTGACACTTACCGGCGCTTCAGGGCACATGGGGGTCTGTCC	2400
QY	2601	TCCCTCTCCCAACCCACCATATGTCTCCAGGGCCCCACTGCAAAAGTTGGCCACTGAGAGCC	2660
Db	2401	TCCCTCTCCCAACCCACCATATGTCTCCAGGGCCCCACTGCAAAAGTTGGCCACTGAGAGCC	2460
QY	2661	ACCTCTTCGCTCCCCCTCTGAGTCTGAGAGTGAACGCTACCTTCTCCCATCCCAACAA	2720
Db	2461	ACCTCTTCGCTCCCCCTCTGAGTCTGAGAGTGAACGCTACCTTCTCCCATCCCAACAA	2520
QY	2721	TGGGGATGTAAAGCAGCAAGACACAGACATTCCTTACTGAACATCAGAGGCCCATGGA	2780
Db	2521	TGGGGATGTAAAGCAGCAAGACACAGACATTCCTTACTGAACATCAGAGGCCCATGGA	2580
QY	2781	GCCAGCAGAAATTAATCTTGATCCATTCACAGGCTTGTGAGTTTCATTAAGCAGAGGCACT	2840
Db	2581	GCCAGCAGAAATTAATCTTGATCCATTCACAGGCTTGTGAGTTTCATTAAGCAGAGGCACT	2640
QY	2841	GAGACACCCGCTCGGTCTCTAACCCAGAAATCTTAAGAAAGAGGAATTTATCAGAAAGAA	2900
Db	2641	GAGACACCCGCTCGGTCTCTAACCCAGAAATCTTAAGAAAGAGGAATTTATCAGAAAGAA	2700
QY	2901	CAGCAGAGAGTTTTCTGTGACACCGCCCACTTCATATGCTCAGTGAATCTTATTCTAAG	2960
Db	2701	CAGCAGAGAGTTTTCTGTGACACCGCCCACTTCATATGCTCAGTGAATCTTATTCTAAG	2760
QY	2961	GCAAGACATTTAAATGATGAATTCCAATCTGATATACATGATACAGCTCATGTGCTCC	3020
Db	2761	GCAAGACATTTAAATGATGAATTCCAATCTGATATACAGCTCATGATCATGTGCTCC	2820
QY	3021	TCAAATTAGGCTGTGCGGTTAGCCAGGCTGTAAAGAGAGAGAGGCTGAGTCACTTA	3080
Db	2821	TCAAATTAGGCTGTGCGGTTAGCCAGGCTGTAAAGAGAGAGAGGCTGAGTCACTTA	2880
QY	3081	GCATAGGGTTGCACAGACGCCCTGTGATTCAGAGTGTAAACAGAGGCTTGCCCTTTAGG	3140
Db	2881	GCATAGGGTTGCACAGACGCCCTGTGATTCAGAGTGTAAACAGAGGCTTGCCCTTTAGG	2940
QY	3141	ACAAACAGTTCCAAATTCAGAGAGCTTACCTGAGTCCCTACTCTGACCTGGGGTCCCAAG	3200
Db	2941	ACAAACAGTTCCAAATTCAGAGAGCTTACCTGAGTCCCTACTCTGACCTGGGGTCCCAAG	3000
QY	3201	ATGAAAAACGAATGTGCTTTTATTAATTAATTTAGTGTGTCTGTGTTATTTAAGA	3260
Db	3001	ATGAAAAACGAATGTGCTTTTATTAATTAATTTAGTGTGTGTGTTATTTAAGA	3060
QY	3261	GATCAAAATGTATTAACAACCTAGCTCTTTTCACTGACTTAAGTATTAATCTACTTAATCTG	3320
Db	3061	GATCAAAATGTATTAACAACCTAGCTCTTTTCACTGACTTAAGTATTAATCTACTTAATCTG	3120
QY	3321	GTTTGGATGCTGGGGTGTGTGACTTCTACTGACCGCTAGATTAACGTGTGCTGTCCCCCA	3380

Db	3121	GTTTGGAAAGCCTGGGGTGTGACTTTCTAATGCAACCGGTAGATAAAGTGCGCTGTCCCCA	3180
Qy	3381	GGTGGTGGGAATAATTTCACAATCTGTCCAAACGAAAAAAGATGTGTGTGTTGAGCAGCA	3440
Db	3181	GGGTGGGAATTAATTTCAAATCTGTCCAAACGAAAAAAGATGTGTGTGTTGAGCAGCA	3240
Qy	3441	TTGACACATATCTGGTGTGATAGAGACTTCCTGTATTCCTAGGCGGTGTGTGTATTC	3500
Db	3241	TTGACACATATCTGTGTTGATAGAGACTTCCTGTATTCCTAGGCGGTGTGTGTATTC	3300
Qy	3501	CCATTGTGAAATTCATCTTGAATCCCATGTGCTTATGTCCTAGCATAAGAGAAATTT	3560
Db	3301	CCATTGTGAAATTCATCTTGAATCCCATGTGCTTATGTCCTAGCATAAGAGAAATTT	3360
Qy	3561	CCTCAAGTTTCCATGTGGCGTTCTCTCTAGCTGCAGCAATACTTGGACATTTAAAGAAA	3620
Db	3361	CCTCAAGTTTCCATGTGGCGTTCTCTCTAGCTGCAGCAATACTTGGACATTTAAAGAAA	3420
Qy	3621	TTTGAAGAAATTTCTCATCCCTCTAAAAATGTTTAAATATACCAAAACAGGGCCCCGTG	3680
Db	3421	TTTGAAGAAATTTCTCATCCCTCTAAAAATGTTTAAATATACCAAAACAGGGCCCCGTG	3480
Qy	3681	CATTAGTTTTCTGTGCGCACTGCACAACCCATTACTTGGAGCTTTAAAAACAACATTAGC	3740
Db	3481	CATTAGTTTTCTGTGCGCACTGCACAACCCATTACTTGGAGCTTTAAAAACAACATTAGC	3540
Qy	3741	TTTATATGTCCTGGGAGATCAGAAATTCGCAAAATGATATGCTCCGTAATGAAATCAAGGTGCA	3800
Db	3541	TTTATATGTCCTGGGAGATCAGAAATTCGCAAAATGATATGCTCCGTAATGAAATCAAGGTGCA	3600
Qy	3801	GCAAGAGCTGTGCTCTCTCTGAAAGGCTCTAGGGAGAAAGCGGTTCTCTGCCATTTCAAGCT	3860
Db	3601	GCAAGAGCTGTGCTCTCTCTGAAAGGCTCTAGGGAGAAAGCGGTTCTCTGCCATTTCAAGCT	3660
Qy	3861	TCTAGAGGCTGGCTGCATTTCCAGAGCTCAGATGGCTCTCAAGCTTTTCTCACATGGCAT	3920
Db	3661	TCTAGAGGCTGGCTGCATTTCCAGAGCTCAGATGGCTCTCAAGCTTTTCTCACATGGCAT	3720
Qy	3921	CAGTGTGACACTGGCCCTCCCACTTCCCTCTTTGACTTACAAAGCCCAACGAGAATCC	3980
Db	3721	CAGTGTGACACTGGCCCTCCCACTTCCCTCTTTGACTTACAAAGCCCAACGAGAATCC	3780
Qy	3981	AGGATTAATCTCTCCATCTTAAAGATCTTTCATCATCTCTGGAAGAGCTTTTGCCATGCAAG	4040
Db	3781	AGGATTAATCTCTCCATCTTAAAGATCTTTCATCATCTCTGGAAGAGAGCTTTTGCCATGCAAG	3840
Qy	4041	ACAACTATGACACAGGTGGGAGATTAGACACAGGACATTTTGGGGGTGTGTTATTCGCC	4100
Db	3841	ACAACTATGACACAGGTGGGAGATTAGACACAGGACATTTTGGGGGTGTGTTATTCGCC	3900
Qy	4101	TACCAACACTTCTCTGCACATGACTCCACAGAGAGGCTACAAATATGATCTGGCGCACAG	4160
Db	3901	TACCAACACTTCTCTGCACATGACTCCACAGAGAGGCTACAAATATGATCTGGCGCACAG	3960
Qy	4161	GGATGTTTGTGTTAGCTTGGCGGACTCTTACACTTAAAAAAAACCAGATCAGGAATCTG	4220
Db	3961	GGATGTTTGTGTTAGCTTGGCGGACTCTTACACTTAAAAAAAACCAGATCAGGAATCTG	4020
Qy	4221	GGCATGCGGGGGCTCACATCTCTACACTAGCAAACTGGCTGAGGCTGGGACCAAGCTCT	4280
Db	4021	GGCATGCGGGGGCTCACATCTCTACACTAGCAAACTGGCTGAGGCTGGGACCAAGCTCT	4080
Qy	4281	GCCATTAGAAAGGGGTGTCTCACTTACAGAGGTCAACACAGCCCACTAAGCCCTATCACT	4340
Db	4081	GCCATTAGAAAGGGGTGTCTCACTTACAGAGGTCAACACAGCCCACTAAGCCCTATCACT	4140
Qy	4341	TCCCAACAATGAGGCTTAAGTGTGTTTCTACTGATCAATGCCCCCTGACAGGTGTGATTTAT	4400
Db	4141	TCCCAACAATGAGGCTTAAGTGTGTTTCTACTGATCAATGCCCCCTGACAGGTGTGATTTAT	4200
Qy	4401	TGTAAATGAAAAAGAAACCTGGGATTAATCTCTTAATCAGGTGATTAACCTATGACCAA	4460
Db	4201	TGTAAATGAAAAAGAAACCTGGGATTAATCTCTTAATCAGGTGATTAACCTATGACCAA	4260

[illegible]

QY	5541	TCGTTTCTTTAAAGCCGAGCCCACTACATTAATAAACTTTACAGGTCACCTGAGAAACAGT	5600
Db	5341	TCGTTTCTTTAAAGCCGAGCCCACTACATTAATAAACTTTACAGGTCACCTGAGAAACAGT	5400
QY	5601	GAAGTGCCATTGTTGAAAGCCCTACCTGACATGCCAGCCCACTGCTACATCCACGCTGGTCTGCC	5660
Db	5401	GAAGTGCCATTGTTGAAAGCCCTACCTGACATGCCAGCCCACTGCTACATCCACGCTGGTCTGCC	5460
QY	5661	ATGCTCTAGAGAAAGGCCACGGCCATTCGACAGACTGGTCTCTAATGCTGTGGTCATTGGACA	5720
Db	5461	ATGCTCTAGAGAAAGGCCACGGCCATTCGACAGACTGGTCTCTAATGCTGTGGTCATTGGACA	5520
QY	5721	GAAGGAAAGAGGCTCAAGGAAAGATGCACCTGGGACCAAGCAAGCCCAACGGACATGGAC	5780
Db	5521	GAAGGAAAGAGGCTCTCAGAGAAAGATGCACCTGGGACCAAGCAAGCCCAACGGACATGGAC	5580
QY	5781	TTGGTAAAGGTTAGACAGACTGGTGTGTGTGAGATTCGACATGCTTCACTGAAATTAATTTA	5840
Db	5581	TTGGTAAAGGTTAGACAGACTGGTGTGTGTGAGATTCGACATGCTTCACTGAAATTAATTTA	5640
QY	5841	TTCAATGCAATACCTTTTAAAGTGCATTTTATTCATTTCTGTGCTTTAATAATAACAA	5900
Db	5641	TTCAATGCAATACCTTTTAAAGTGCATTTTATTCATTTCTGTGCTTTAATAATAACAA	5700
QY	5901	TGTACCAAAAAACAAGTATCAAGCTGTTAAGTGTGTGGCTACTGTGCTCCCTGGTTCAG	5960
Db	5701	TGTACCAAAAAACAAGTATCAAGCTGTTAAGTGTGTGGCTACTGTGCTCCCTGGTTCAG	5760
QY	5961	TAGAGAGCCCGCGTGTCCAGTTGTGTGACTGTGACAGGCTCAGACATGGGCTCAGACAGATGC	6020
Db	5761	TAGAGAGCCCGCGTGTCCAGTTGTGTGACTGTGTGACAGGCTCAGACATGGGCTCAGACAGATGC	5820
QY	6021	TGCTCTAATTTGTGTGATGATACAGAAAGCCAGGCTTTGGATACAAAGTCTTTCTCTTC	6080
Db	5821	TGCTCTAATTTGTGTGATGATACAGAAAGCCAGGCTTTGGATACAAAGTCTTTCTCTTC	5880
QY	6081	ATTGTAGCCCGTGACCTGTGTGAAGACAAATGTTTTGTCCGAAATTAATAATTAATAGCT	6140
Db	5881	ATTGTAGCCCGTGACCTGTGTGAAGACAAATGTTTTGTCCGAAATTAATAATTAATAGCT	5940
QY	6141	TGGAGTCTGGCCAAAAAATAAAAAA 6163	
Db	5941	TGGAGTCTGGCCAAAAAATAAAAAA 5963	
RESULT 3			
LOCUS	AX353603	5897 bp	DNA linear PAT 06-FEB-2002
DEFINITION	Sequence 1 from Patent. WO0204508.		
ACCESSION	AX353603		
VERSION	AX353603.1	GI:18618676	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE			
AUTHORS	1		
TITLE	Schweifer,N., Scheerl-Mostaguer,M., Sommergruber,W., and Abseher,R.		
FEATURES			
Journal	Tumour-associated antigen (b345), characterised by an amino acid		
Patent	sequence as in seq. Id. No. 4		
Location	MO 0204508-A 1 17-JAN-2002;		
Qualifiers			
1.	5897		
2.	5897		
3.	5897		
4.	5897		
5.	5897		
6.	5897		
7.	5897		
8.	5897		
9.	5897		
10.	5897		
11.	5897		
12.	5897		
13.	5897		
14.	5897		
15.	5897		
16.	5897		
17.	5897		
18.	5897		
19.	5897		
20.	5897		
21.	5897		
22.	5897		
23.	5897		
24.	5897		
25.	5897		
26.	5897		
27.	5897		
28.	5897		
29.	5897		
30.	5897		
31.	5897		
32.	5897		
33.	5897		
34.	5897		
35.	5897		
36.	5897		
37.	5897		
38.	5897		
39.	5897		
40.	5897		
41.	5897		
42.	5897		
43.	5897		
44.	5897		
45.	5897		
46.	5897		
47.	5897		
48.	5897		
49.	5897		
50.	5897		
51.	5897		
52.	5897		
53.	5897		
54.	5897		
55.	5897		
56.	5897		
57.	5897		
58.	5897		
59.	5897		
60.	5897		
61.	5897		
62.	5897		
63.	5897		
64.	5897		
65.	5897		
66.	5897		
67.	5897		
68.	5897		
69.	5897		
70.	5897		
71.	5897		
72.	5897		
73.	5897		
74.	5897		
75.	5897		
76.	5897		
77.	5897		

3' UTR
BASE COUNT
ORIGIN

1464 a 1596 c 1364 g 1467 t 6 others

2465 .5897
PSDIPSTMGM⁺
IPAPSCOSQRMVITGRSRAPWGSVLPHPPIADGPOLQSWPLRSHLALPLMLRNRT

YKKKKKKTKNGPAVGIYNGNINTEMPGSSQSPKRGERTPMQMOQSRTYMWNGICVR
HHSFWNINSNCSTSGKODLPSVTLTPGTVLIIYLAAGGVLLSALGILICCC
SNMISVPRDVALCTFFKRSQVAVQCTGAFMIIOERTAAEIBFSLDEVLVLPSPFH
LIRFAPSPQEASRQGLTVSFIYFEKEEGVFTVPTKSAVYLRTPMWDGLPSLTVS
LUVPRKLOQOHTHEKPCNTSFYLYASALPSODLITFGSCPCPGSITKOIYQONISV
CDLFTLMMNVKETISCTDHRVCCRSKYSALPSVDLIILHPEVLDHPSWKLIVPKDRIS
KIYVLDINSERAMSLITIEPRPYQSKRPFGSCVCSJESRPSGNLITLSSKKIKSFL
IPSTNNIEVFYKLEDPQNGMAGNFNLSLOGCDODASPGILIEIQFVLAQHPQNSN
PFBESGATLMSANYPGEFPEDELMTQFVAPALURASVSLNINLSCKEKEVEEYV
ATVYRIGTPELPCNTSVRIKQOEGVKNMALLPWHPRNPNVSGFISANSIRGLCTISB

LOPSTLPLTNTPIIMDVKAHSIGLELOFISIPRLRQIGFSSGCPGVNTHISGRID

Query Match	87.5%;	Score 5395;	DB 6;	Length 5897;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 5865;	Conservative	0;	Mismatches 7;	Indels 1;
				Gaps 1;
QY	291	CTTGAACCTGCGGGGTCCTATATGCACTCTGTAAGGGGTTCTGCTGCTGGGTGCGCGCCCT	350	
Db	25	CTTGAACCTGCGGGGTCCTATATGCACTCTGTAAGGGGTTCTGCTGCTGGGTGCGCGCCCT	84	
QY	351	GCGCGCGGGGCGAAGACCTTTTGAATGCTCTGCGCCACGAGAAAGCAACTTACAGTTCT	410	
Db	85	GCGCGCGGGGCGAAGACCTTTTGAATGCTCTGCGCCACGAGAAAGCAACTTACAGTTCT	144	
QY	411	CATAAAGCTGGGAGACCCCGACCTGCTGCGCAAAACCTGTACATCGTCACTTTCTAAAG	470	
Db	145	CATTAAGCTGGGAGACCCCGACCTGCTGCGCAAAACCTGTATCATGTCATTTCTAAAG	204	
QY	471	ACATATTAACATGTTGTTCATCAAGTCTGGAAGAAATGTCTTTACCTTTAGCTGCCA	530	
Db	205	ACATATTAACATGTTGTTCATCAAGTCTGGAAGAAATGTCTTTACCTTTAGCTGCCA	264	
QY	531	GAGGCCGAGAATCACTTTGTCATAGAGATCCAGAAATAATGACGTATGTGAGGCC	590	
Db	265	GAGTCTCGAATATCATTTTGTCTATAGAGATCCAGAAATAATGACGTATGTGAGGCC	324	
QY	591	ATGTCCTTTTGGGAGGTTCAAGCTTCAAGCCCTGACATCGTTGTGCTTACCCCTCAAG	650	
Db	325	ATGTCCTTTTGGGAGGTTCAAGCTTCAAGCCCTGACATCGTTGTGCTTACCCCTCAAG	384	
QY	651	AACCTTCATCTGGGATGTCAAACCTCATTAAGACATCGTTTGAAGTGCAGATTTTCAT	710	
Db	385	AACCTTCATCTGGGATGTCAAACCTCATTAAGACATCGTTTGAAGTGCAGATTTTCAT	444	
QY	711	CCCTCGCTGAGGCAAGTCGGTCCGGGTGAGACTGCCCGAGCGAGTCACTCACTCAT	770	
Db	445	CCCTCGCTGAGGCAAGTCGGTCCGGGTGAGACTGCCCGAGCGAGTCACTCACTCAT	504	
QY	771	CAGCGCGCAATCATGCCACCGTGTCAAGATCGAAACCTTCTGCAATGGCACTGT	830	
Db	505	CAGCGCGCAATCATGCCACCGTGTCAAGATCGAAACCTTCTGCAATGGCACTGT	564	
QY	831	GTCGCCGATCAAGATGCAAGAGAGTGAATAGCCTTACACTCCACATGTTCCACC	890	
Db	565	GTCGCCGATCAAGATGCAAGAGAGTGAATAGCCTTACACTCCACATGTTCCACC	624	
QY	891	CAGAAATGTCTCCGGCTTCAGATTGCAAAACCGCTCATATATAAAAGTCTGTGCATCAT	950	
Db	625	CAGAAATGTCTCCGGCTTCAGATTGCAAAACCGCTCATATATAAAAGTCTGTGCATCAT	684	
QY	951	CGAGTCTGTGTTTGAAGGTGAAGGCTTCAGCAACCTGTATGTCTGCCAACTACCCAGAG	1010	
Db	685	CGAGTCTGTGTTTGAAGGTGAAGGCTTCAGCAACCTGTATGTCTGCCAACTACCCAGAG	744	
QY	1011	CTTCCCGAGAGATGAGCTCATAGACGTGCGACGTTTGTCTCCGCAACCTGGGGGCGAG	1070	
Db	745	CTTCCCGAGAGATGAGCTCATAGACGTGCGACGTTTGTCTCCGCAACCTGGGGGCGAG	804	

QY	1071	GGTCTCCCTTCCCTCAACTTCAACCTCTCCAACTGTGAGAGGAAGAGACGGGTGTGAATA	1130
Db	805	CGTCTCCCTTCCCTCAACTTCAACCTCTCCAACTGTGAGAGGAAGAGACGGGTGTGAATA	864
QY	1131	CTACATCCCGGGGCTCCACCAACCAACCCCGAGGTGTTCAAGCTGGAGGACAAGACGCTGG	1190
Db	865	CTACATCCCGGGGCTCCACCAACCAACCCCGAGGTGTTCAAGCTGGAGGACAAGACGCTGG	924
QY	1191	GAACATGGCGGGGGAATTCAACCTCTCTGCAAGGCTGTGACCAAGATGCCCCAAAGTCC	1250
Db	925	GAACATGGCGGGGGAATTCAACCTCTCTGCAAGGCTGTGACCAAGATGCCCCAAAGTCC	984
QY	1251	AGGGAATCCTCGGGCTGCGATTCCAAGTTTGGTCCAAATCCACAAAATGAAGCATTA	1310
Db	985	AGGGAATCCTCGGGCTGCGATTCCAAGTTTGGTCCAAATCCACAAAATGAAGCATTA	1044
QY	1311	AATCTACGTGGTTGACTTGAGTAATGAGCGAGCATGCTCAACATCGAGCCACGGCC	1370
Db	1045	AATCTACGTGGTTGACTTGAGTAATGAGCGAGCATGCTCAACATCGAGCCACGGCC	1104
QY	1371	CGTCAAAACAGACCCGCAAGTTTGTCCCTGGCTGTTCCGTGTGTCTAGAAATCTCGAACCTG	1430
Db	1105	CGTCAAAACAGACCCGCAAGTTTGTCCCTGGCTGTTCCGTGTGTCTAGAAATCTCGAACCTG	1164
QY	1431	CAGTAGCAACTTCACTCACTCGATCATCTGGCTCCAAACACAAATCTCTCTTTGTATGA	1490
Db	1165	CAGTAGCAACTTCACTCACTCGATCATCTGGCTCCAAACACAAATCTCTCTTTGTATGA	1224
QY	1491	TCTGACACGCTGTGTGATGAATGTGTGAAAAACAATAAGCTGCACAGACACCGGTTACTG	1550
Db	1225	TCTGACACGCTGTGTGATGAATGTGTGAAAAACAATAAGCTGCACAGACACCGGTTACTG	1284
QY	1551	CCAAAGGAATCTTACTCACTCCAGGTGCCCCAGTGCACATCTCTCGACCTGCTGTGAGCT	1610
Db	1285	CCAAAGGAATCTTACTCACTCCAGGTGCCCCAGTGCACATCTCTCGACCTGCTGTGAGCT	1344
QY	1611	GCATGACTTCTCTCTGGAAAGCTGTGTGTGCCAAGGACAGGCTCAGGCTGTGTGTGTGCC	1670
Db	1345	GCATGACTTCTCTCTGGAAAGCTGTGTGTGCCAAGGACAGGCTCAGGCTGTGTGTGTGCC	1404
QY	1671	AGCCCAGAGCTGGAGGACGATTAACACGAAAGCCCTGGCAACACGAGCTTCACTCACT	1730
Db	1405	AGCCCAGAGCTGGAGGACGATTAACACGAAAGCCCTGGCAACACGAGCTTCACTCACT	1464
QY	1731	CGTGGCCAGTGCCTAATCCCAACCCAGGACCTGTACTTCCGCTCTCTTCTGCCCCGAGAGCTC	1790
Db	1465	CGTGGCCAGTGCCTAATCCCAACCCAGGACCTGTACTTCCGCTCTCTTCTGCCCCGAGAGCTC	1524
QY	1791	TATCAAGCAGATCCAGGTGAGGACGAAACATCTCGGTGAACCTCTTGCGACCTTTTGCCCCAG	1850
Db	1525	TATCAAGCAGATCCAGGTGAGGACGAAACATCTCGGTGAACCTCTTGCGACCTTTTGCCCCAG	1584
QY	1851	CTTCCCAACAAAGGCTCTCCAGGACAGGCTCTACCGGTGCTCTTTATACCTTAATTCCAAAGA	1910
Db	1585	CTTCCCAACAAAGGCTCTCCAGGACAGGCTCTACCGGTGCTCTTTATACCTTAATTCCAAAGA	1644
QY	1911	GGAAGGCGTTTTCACGAGTACCCCTGACACAAAAGCAAGGCTTAACCTGAGGACCCCCCAA	1970
Db	1645	GGAAGGCGTTTTCACGAGTACCCCTGACACAAAAGCAAGGCTTAACCTGAGGACCCCCCAA	1704
QY	1971	CTGGGACCGGGGGCTGCGCATCTTCAACCTCTGTGTCTCTGGAACATCAGCGTGCACAGAGA	2030
Db	1705	CTGGGACCGGGGGCTGCGCATCTTCAACCTCTGTGTCTCTGGAACATCAGCGTGCACAGAGA	1764
QY	2031	CCAGGTGGCTTGTCTGACTTTCTTTAAGAGACGGAACGCGCTGGTCTGTGCCAGACAGGGCG	2090
Db	1765	CCAGGTGGCTTGTCTGACTTTCTTTAAGAGACGGAACGCGCTGGTCTGTGCCAGACAGGGCG	1824
QY	2091	CGCAATTCAATGATTCACAGAGACGCGGACCCCGGGCTGAGAGATCTTCAACCTGTGAGAGA	2150
Db	1825	CGCAATTCAATGATTCACAGAGACGCGGACCCCGGGCTGAGAGATCTTCAACCTGTGAGAGA	1884
QY	2151	GGATGTGCTCCCAAGCCAGCTTCCACATCAACGCTTGTGGGTCAACATCTCTAACTG	2210

Db	1885	GGATGTGCTCCCAAGCCAGCTTCCACATCACAGCTTCTGGGTCAACATCTCTAACTG	1944
QY	2211	CAGCCCAACGACCGGCAAGCAGTACGACTGCTCTTCTCGGTGACATTACCCCAAGAC	2270
Db	1945	CAGCCCAACGACCGGCAAGCAGCTACGACTGCTCTTCTCGGTGACATTACCCCAAGAC	2004
QY	2271	TGTGGACTTGACTGATCCTCATCGACCGGTGGAGGTGAGTCTTACTGCTGTCCG	2330
Db	2005	TGTGGACTTGACTGATCCTCATCGACCGGTGGAGGTGAGTCTTACTGCTGTCCG	2064
QY	2331	CCTGGGGCTCATCTTTGCTGTGTGAAAAAGAAAGAAAGAAAGAAAGAGGCCCCGC	2390
Db	2065	CCTGGGGCTCATCTTTGCTGTGTGAAAAAGAAAGAAAGAAAGAGGCCCCGC	2124
QY	2391	TGTGGGTATCTACATGGCAATCAATCAATCTGAGATCCGAGGCAAGCCAAAGTTTCA	2450
Db	2125	TGTGGGTATCTACATGGCAATCAATCAATCTGAGATCCGAGGCAAGCCAAAGTTTCA	2183
QY	2451	GAAGGGCGAAGGACAAATGACTCCCATGTGTATGACATCATGAGACACCATGTATTA	2510
Db	2184	GAAGGGCGAAGGACAAATGACTCCCATGTGTATGACATCATGAGACACCATGTATTA	2243
QY	2511	TGGGCATCTGTAAAGATTCACGGGCTCCTTCCTGACGACGAGGTGACACTTACCG	2570
Db	2244	TGGGCATCTGTAAAGATTCACGGGCTCCTTCCTGACGACGAGGTGACACTTACCG	2303
QY	2571	GCCTGTCAGGGGACCAATGAGGGGTCTGTCCCTCCGCCCAACCCCAATATGCTCAAGGC	2630
Db	2304	GCCTGTCAGGGGACCAATGAGGGGTCTGTCCCTCCGCCCAACCAATATGCTCAAGGC	2363
QY	2631	CCCAACTGCAAGTGGGCACTGAGAGCCACTCTCGCTCCCTCTGAGTGTAGAG	2690
Db	2364	CCCAACTGCAAGTGGGCACTGAGAGCCACTCTCGCTCCCTCTGAGTGTAGAG	2423
QY	2691	TGAACCGTACCTTTCTCCATCCCAACATGGGGATGTAAAGCAGACAGACAT	2750
Db	2424	TGAACCGTACCTTTCTCCATCCCAACATGGGGATGTAAAGCAGACAGACAT	2483
QY	2751	TCCCTTACTGAACCTCAGAGCCCATGAGCCAGAGAAATTAATTGATCATTTCCAGAC	2810
Db	2484	TCCCTTACTGAACCTCAGAGCCCATGAGCCAGAGAAATTAATTGATCATTTCCAGAC	2543
QY	2811	GCTTGTGTAATTCATTAAGCAGGGGCACTGAGACACCCGTCGTCGTTCTTAACAGAAA	2870
Db	2544	GCTTGTGTAATTCATTAAGCAGGGGCACTGAGACACCCGTCGTCGTTCTTAACAGAAA	2603
QY	2871	TCTTAAGAGAGAAATTATACAGAGGAAACAGCAGAGGTTTTCTGAGACCCGCCAAC	2930
Db	2604	TCTTAAGAGAGAAATTATACAGAGGAAACAGCAGAGGTTTTCTGAGACCCGCCAAC	2663
QY	2931	TTTCACTTTGCTCAGTGAATCATTTCTTAAGGGCAAGATTTGAAAAATATGATTTCCATC	2990
Db	2664	TTTCACTTTGCTCAGTGAATCATTTCTTAAGGGCAAGATTTGAAAAATATGATTTCCATC	2723
QY	2991	TGATATACGTCATACAGCTCATGTGCTCTCAACTTAAGCTGTGCGGTTAGCCAGCTG	3050
Db	2724	TGATATACGTCATACAGCTCATGTGCTCTCAACTTAAGCTGTGCGGTTAGCCAGCTG	2783
QY	3051	TAAATGAGAGAGAGGCTGATGCTACCTTAAGCATAGGGTTGACAGACCCCTGGATTGAG	3110
Db	2784	TAAATGAGAGAGAGGCTGATGCTACCTTAAGCATAGGGTTGACAGACCCCTGGATTGAG	2843
QY	3111	AGTGTAAACAGAGGCTTGCCCTTTCAGAGCAACATTTCCAAATTCAGAGAGCTTACT	3170
Db	2844	AGTGTAAACAGAGGCTTGCCCTTTCAGAGCAACATTTCCAAATTCAGAGAGCTTACT	2903
QY	3171	GAGGTCCCTACTCTCATCTGGGGTCCCAAGATGAAACGACAAATGTGCTTTTATTAATT	3230
Db	2904	GAGGTCCCTACTCTCATCTGGGGTCCCAAGATGAAACGACAAATGTGCTTTTATTAATT	2963
QY	3231	ATTATATTGTGTGCTGTGTATTTAAGATCAAAATGTATTAACCACTTACTCTTTTC	3290

DB 2964 ATTATTTGGTGTCTCTGTCTTATTTAAGAGATCAAAATGATTAACCACTAGCTTTTTC 3023
QY 3291 ACCTGACTTATGATTAATCATATCATACTGTTGGATGCTGGTTGACTTCTACTG 3350
DB 3024 ACCTGACTTATGATTAATCATATCATACTGTTGGATGCTGGTTGACTTCTACTG 3083
QY 3351 ACCGCTAGATTAACGTTGCTCTGCTCCCAAGTGTGGGAATATTTAACAATCTGTCAA 3410
DB 3084 ACCGCTAGATTAACGTTGCTCTGCTCCCAAGTGTGGGAATATTTAACAATCTGTCAA 3143
QY 3411 CCGAAGAAAGATGTGTGTGTGTAGAGAGATGACATATCTGCTTTGATTAAGAGACT 3470
DB 3144 CCGAAGAAAGATGTGTGTGTGTAGAGAGATGACATATCTGCTTTGATTAAGAGACT 3203
QY 3471 CCTGATTTCTTAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3530
DB 3204 CCGATTTCTTAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3263
QY 3531 GTCTCTATGCTCTAGCAATAGAGAAATTTCTCAAGTTTCCATGTGCGGTTCTCTAGC 3590
DB 3264 GTCTCTATGCTCTAGCAATAGAGAAATTTCTCAAGTTTCCATGTGCGGTTCTCTAGC 3323
QY 3591 TGAGAGAAATCTTGACATTTAAGAGAAATTTAGAGAAATCTGATCTCTAAGAAATG 3650
DB 3324 TGAGAGAAATCTTGACATTTAAGAGAAATTTAGAGAAATCTGATCTCTAAGAAATG 3383
QY 3651 TTTAAATATATACCAACAGTGGCCCCCTGATTTGTTTCTGTTGCTGCTGCTGCTGCT 3710
DB 3384 TTTAAATATATACCAACAGTGGCCCCCTGATTTGTTTCTGTTGCTGCTGCTGCTGCT 3443
QY 3711 TACTGTGATCTTAAACCAACATTTAGCTTATGCTGCTGGGATCGAATTTCCAAAT 3770
DB 3444 TACTGTGATCTTAAACCAACATTTAGCTTATGCTGCTGGGATCGAATTTCCAAAT 3503
QY 3771 GATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3830
DB 3504 GATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3563
QY 3831 GGAGAGCCGGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3890
DB 3564 GGAGAGCCGGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3623
QY 3891 GTGGCTGTGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3950
DB 3624 GTGGCTGTGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3683
QY 3951 TTTGATCTTCAAGCCCAACAGAGATTCAGAGATTCCTGATCTAAGATCTCTCA 4010
DB 3684 TTTGATCTTCAAGCCCAACAGAGATTCAGAGATTCCTGATCTAAGATCTCTCA 3743
QY 4011 TCATCTCTGAGAGAGCTTTTGCATGCAAGACATAGCCACAGGTGGGATTTAGAGCC 4070
DB 3744 TCATCTCTGAGAGAGCTTTTGCATGCAAGACATAGCCACAGGTGGGATTTAGAGCC 3803
QY 4071 AGGACATCTTTGGGGGTCTGTTATTTGCTTACCAACCTTCTGCACTGACTCCACA 4130
DB 3804 AGGACATCTTTGGGGGTCTGTTATTTGCTTACCAACCTTCTGCACTGACTCCACA 3863
QY 4131 GGAGAGGCTTCAAAATGATCTGGGCAAGAGAGATTTGTTTGAAGCTTGGGACTTAAC 4190
DB 3864 GGAGAGGCTTCAAAATGATCTGGGCAAGAGAGATTTGTTTGAAGCTTGGGACTTAAC 3923
QY 4191 ACTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4250
DB 3924 ACTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3983
QY 4251 AACCACTGCTGAGCTGGGCAACAGCTGCTTTTGAAGGGGTCTCACTTCAACAG 4310
DB 3984 AACCACTGCTGAGCTGGGCAACAGCTGCTTTTGAAGGGGTCTCACTTCAACAG 4043
QY 4311 TCACCAAGCCCACTAGCCCTTATCTTCCCAATGAGGCTAAGTGTGTTCTA 4370
DB 4044 TCACCAAGCCCACTAGCCCTTATCTTCCCAATGAGGCTAAGTGTGTTCTA 4103

QY 4371 CTGATCAATGCCCTGCAAGTTGATTTATTTGTAATGAAGAAAGAACTGGATTAATC 4430
DB 4104 CTGATCAATGCCCTGCAAGTTGATTTATTTGTAATGAAGAAAGAACTGGATTAATC 4163
QY 4431 TCTATTCAGTGAAGTGAACATAGAACCAATGTGTGCTCACTTACCTTTCTTTT 4490
DB 4164 TCTATTCAGTGAAGTGAACATAGAACCAATGTGTGCTCACTTACCTTTCTTTT 4223
QY 4491 TCTATTCAGTGAAGTGAACATAGAACCAATGTGTGCTCACTTACCTTTCTTTT 4550
DB 4224 TCTATTCAGTGAAGTGAACATAGAACCAATGTGTGCTCACTTACCTTTCTTTT 4283
QY 4551 TGAAGTGAAGTGGCAATCTGCTCACTGCAACCTGCTCTGCTGCTCAAGCAAT 4610
DB 4284 TGAAGTGAAGTGGCAATCTGCTCACTGCAACCTGCTCTGCTGCTCAAGCAAT 4343
QY 4611 CTCCCACTGAGCTTCCCAAAATGCTGGGATTCAGTGGCAAAACCACTGCTGCTTAA 4670
DB 4344 CTCCCACTGAGCTTCCCAAAATGCTGGGATTCAGTGGCAAAACCACTGCTGCTTAA 4403
QY 4671 TTTGATTTTGTAGAGACAGGTTTCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCT 4730
DB 4404 TTTGATTTTGTAGAGACAGGTTTCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCT 4463
QY 4731 GGGCTCAAGCAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4790
DB 4464 GGGCTCAAGCAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4523
QY 4791 CCGATCCAGCCCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4850
DB 4524 CCGATCCAGCCCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4583
QY 4851 GCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4910
DB 4584 GCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4643
QY 4911 GTCCCAAGTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4970
DB 4644 GTCCCAAGTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4703
QY 4971 TATTCATGATTTCACTTTGGAATGCAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5030
DB 4704 TATTCATGATTTCACTTTGGAATGCAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4763
QY 5031 AGGAATTTTATGCTCTCTTAAATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5090
DB 4764 AGGAATTTTATGCTCTCTTAAATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4823
QY 5091 GATAGGATTTTATGCTCTCTTAAATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5150
DB 4824 GATAGGATTTTATGCTCTCTTAAATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4883
QY 5151 TGGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5210
DB 4884 TGGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4943
QY 5211 CAAAGAGGGCTGCGGCTTCTTCTCTCACTGAAGAGCCCTTATTTGAATTAATCACTGTGTG 5270
DB 4944 CAAAGAGGGCTGCGGCTTCTTCTCTCACTGAAGAGCCCTTATTTGAATTAATCACTGTGTG 5003
QY 5271 GAGCCCTAGCCCTGCAATCTGCAATTTCCCAACCTTCCCAAGCCCTTCCCAAGCAGACTA 5330
DB 5004 GAGCCCTAGCCCTGCAATCTGCAATTTCCCAACCTTCCCAAGCCCTTCCCAAGCAGACTA 5063
QY 5331 GGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5390
DB 5064 GGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5123
QY 5391 TCACCAAGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5450
DB 5124 TCACCAAGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5183

QY 5451 TCAAACTCTGGAGCAAGTGTCTGCTCTGAGACCAAGTATTTCTGAGCTGTGCTCA 5510
DB 5184 TCAAACTCTCTGGAGCAAGTGTCTGCTCTGAGACCAAGTATTTCTGAGCTGTGCTCA 5243
QY 5511 GTGAAGGGGGCCAGCTGAGGAACCTGGCTCTTTTCTTAAAGCCAGCCCACTTAC 5570
DB 5244 GTGAAGGGGGCCAGCTGAGGAACCTGGCTCTTTTCTTAAAGCCAGCCCACTTAC 5303
QY 5571 ATAAAACATTTGAGGCTCACTGGAACCAAGTGAAGTCCATTTGTTGAAGCTCACTGATG 5630
DB 5304 ATAAAACATTTGAGGCTCACTGGAACCAAGTGAAGTCCATTTGTTGAAGCTCACTGATG 5363
QY 5631 CCAGCCCACTGCTCATCCAGCTGTCTGCTGATGCTTACGAGAAAGCCAGCCCACTGATG 5690
DB 5364 CCAGCCCACTGCTCATCCAGCTGTCTGCTGATGCTTACGAGAAAGCCAGCCCACTGATG 5423
QY 5691 ACTGCTCTTAATGCTGTGTGTGCTGATGCTTACGAGAAAGCTCTCAAGAAAGTCACT 5750
DB 5424 ACTGCTCTTAATGCTGTGTGTGCTGATGCTTACGAGAAAGCTCTCAAGAAAGTCACT 5483
QY 5751 GGGAGACAGCAAGCCCAAGCCAGCAATGCTGTGTAAGGTTAGCAGACTGGTGTGTG 5810
DB 5484 GGGAGACAGCAAGCCCAAGCCAGCAATGCTGTGTAAGGTTAGCAGACTGGTGTGTG 5543
QY 5811 GATTCGACGTGCTTCACTGGAATAATTTATTCATTCAGATACCTTTTAAAGTGGCATTT 5870
DB 5544 GATTCGACGTGCTTCACTGGAATAATTTATTCATTCAGATACCTTTTAAAGTGGCATTT 5603
QY 5871 TATTCATTTCTGCTGCTTTAATTAACCAATGTACCAAAAAACAGTATCAAGCTGTTTA 5930
DB 5604 TATTCATTTCTGCTGCTTTAATTAACCAATGTACCAAAAAACAGTATCAAGCTGTTTA 5663
QY 5931 AGTGTCTGCGCTACTTGTCTCCCTGCTGTGAGAGGCCCCGGTTTCCAGTGTGTGACTG 5990
DB 5664 AGTGTCTGCGCTACTTGTCTCCCTGCTGTGAGAGGCCCCGGTTTCCAGTGTGTGACTG 5723
QY 5991 TGACAGGCTCAGCATGAGGCTCAGCAGATGCTGTAAATTTGTGATGATACAGAAAGCC 6050
DB 5724 TGACAGGCTCAGCATGAGGCTCAGCAGATGCTGTAAATTTGTGATGATACAGAAAGCC 5783
QY 6051 AGGCTTTGGGATACAAAGTCTTTCTCTTCATTTGATGCTGCTGATGATGATGATGATG 6110
DB 5784 AGGCTTTGGGATACAAAGTCTTTCTCTTCATTTGATGCTGCTGATGATGATGATGATG 5843
QY 6111 GTTTTGTCCGGAAATTAATAATATAGCTTGGAGCTGCGCAAAAAA 6163
DB 5844 GTTTTGTCCGGAAATTAATAATATAGCTTGGAGCTGCGCAAAAAA 5896

RESULT 4
AF468010 5978 bp mRNA linear PRI 02-FEB-2003
LOCUS AF468010
DEFINITION Homo sapiens NCSG135 mRNA, complete cds.
ACCESSION AF468010.1 GI:28194091
VERSION AF468010.1
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homosapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 5978)
Hooper,J.D., Zijlstra,A., Aimes,R.T., Claassen,G.F. and
Ouilley,J.P.
TITLE Subtractive Immunization Identifies NCSG-135, a Novel Cell Surface
Glycoprotein
JOURNAL Unpublished
AUTHORS Hooper,J.D., Zijlstra,A., Aimes,R.T., Claassen,G.F. and
Ouilley,J.P.
TITLE Direct Submission
JOURNAL Submitted (14-JAN-2002) Cell Biology, Scripps Research Institute,
10550 North Torrey Pines Road, La Jolla, CA 92037, USA
FEATURES Location/Qualifiers

source
1. 5978
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="3"
/map="3p21.32"
/note="novel cell surface glycoprotein 135 kDa"
/codon_start=1
/product="NCSG135"
/protein_id="AA033397.1"
/db_xref="GI:28194092"
/translation="MAGLNCGVSIALLGVLLGAARLPDGAERFALPESNTVLI
KLGPTTLAKPCYIVISKRHITLMSIKGSEIVFPCOSPEMNHVIEIORKIDCMG
PCPGEVOLQSTSLPLNTRFTIMDVANISIGLEIFSLRLEIGPESCPMGV
HSISGRIDATVIRIGTFCNSNGVSRIRKQEKVAKLHPMPHPRVAVSGPILNRSIK
RLCTIESIFPBGSTLMSANYPGEPFDELTWQFVPAHRAVSELTNPLNSCER
KEEREYVTPGSSTNPEVKLEDKQPGMAGNFNLSLQCDQDQSPGILRLQFVLV
QHPONESNKIIVVDLSNERAMSLTTEPRPVQSRKRFVGCFLSRCSNLTLSG
SKHISFLCDBLTRLNANVEKTICTDHRVQKRSYSLQVPSDLHLLEVLHDFMWL
LVPKRLSLVLPVPAOKLOHHEKRCNTSFYLVSALPSODLYGSCFGGSIKOI
VKONISVTLRTFAPSPFROASROGLTVAFTIYKRGVETVTPDTRKSVTLTPMDR
GLPSLTISVSNISVPRDYACTTFPKESGVCCGRAPMIIQERTABEITFSD
VLKPSFHHHSFWNISNCSPTSGKQDLFLSVTLTPRTVLTLLVLAAGVLLLS
ALGLIICVKKKKKTKNGPAGVGINENINTEMPQPKFQGRKDNDSHYAVLEDT
MVGHLDDSSGSFLQPEVDYRPFQGTMGVCPSPPTICSRAPAKLATEPPRSP
PESSEEPYTFSPNNGDVSSKDTDIPILSTQEPMPAP"

BASE COUNT 1457 a 1637 c 1407 g 1477 t
ORIGIN

Query Match 81.1%; Score 496; DB 9; Length 5978;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 5966; Conservative 0; Mismatches 11; Indels 4; Gaps 3;

QY 173 GCGCCAGGTGATGATGACGAGGCGGAGCGAGCTGCGCGGCGCTGGGCGCTGGGCGC 232
DB 1 GCGCCAGGTGATGATGACGAGGCGGAGCGAGCTGCGCGGCGCTGGGCGCTGGGCGC 60
QY 223 GCGGCTCCCAACCGCGTTTCCCAACGAGGCGGAGGCTCGGAGATGAGGCGGCGC 292
DB 61 GCGGCTCCCAACCGCGTTTCCCAACGAGGCGGAGGCTCGGAGATGAGGCGGCGC 120
QY 293 TGAATCGGGGGCTCTATACGACCTGATGAGGGTTCTGCTGCTGGAGTGGGCGGCGCTGC 352
DB 121 TGAATCGGGGGCTCTATACGACCTGATGAGGGTTCTGCTGCTGGAGTGGGCGGCGCTGC 180
QY 353 GCGCGGGGCGAGACCTTTTGAATGATGCTCTGCCACGAGAAAGCAACATTACAGTTCTCA 412
DB 181 GCGCGGGGCGAGACCTTTTGAATGATGCTCTGCCACGAGAAAGCAACATTACAGTTCTCA 240
QY 413 TAAAGCTGGGACCCCGACTGCTGCGCAAAAACCTGTATACCTCTATTCTAATAAGC 472
DB 241 TAAAGCTGGGACCCCGACTGCTGCGCAAAAACCTGTATACCTCTATTCTAATAAGC 300
QY 473 ATATAACATGTTTGCATCAAGTCTGAGAAAGATAGTCTTTACCTTTAGCTCCAGCA 532
DB 301 ATATAACATGTTTGCATCAAGTCTGAGAAAGATAGTCTTTACCTTTAGCTCCAGCA 360
QY 533 GTCTGAGATCATCTTTGTATAGATCCAGAAAAATATGACTGTATGTCAGGCCAT 592
DB 361 GTCTGAGATCATCTTTGTATAGATCCAGAAAAATATGACTGTATGTCAGGCCAT 420
QY 593 GTCTTTTGGGAGGTTCAAGCTTCAAGCCCTGCAATGCTGTGTGGCTAACCTTCACAGAA 652
DB 421 GTCTTTTGGGAGGTTCAAGCTTCAAGCCCTGCAATGCTGTGTGGCTAACCTTCACAGAA 480
QY 653 CTTTCATCTGGAGATCAAAAGCTCATAGAGAGATGAGGTTTGAAGCTGAGTTTTCATCC 712
DB 481 CTTTCATCTGGAGATCAAAAGCTCATAGAGAGATGAGGTTTGAAGCTGAGTTTTCATCC 540
QY 713 CTGCGCTGAGGAGATCGGCTCGGAGTGAAGCTGCCAGACGAGATCACTACCTCATCA 772
DB 541 CTGCGCTGAGGAGATCGGCTCGGAGTGAAGCTGCCAGACGAGATCACTACCTCATCA 600

QY 773 GCGGCGGAATCGATGCCACCGTGTGAGATCGGAACTTTGTGACAAATGGCATTGTGT 832
Db 601 GCGGCGGAATCGATGCCACCGTGTGAGATCGGAACTTTGTGACAAATGGCATTGTGT 660
QY 833 CCGGGATCGAAGATGCAAGAAAGAGTAAAAATGGCTTACACCTCCATAGTGTCCACCCA 892
Db 661 CCGGGATCGAAGATGCAAGAAAGAGTAAAAATGGCTTACACCTCCATAGTGTCCACCCA 720
QY 893 GAAATGTCTCCGGCTTCAGCATTTGCAAAACCGCTCATCTATATAAGCTGTGTGATCATCG 952
Db 721 GAAATGTCTCCGGCTTCAGCATTTGCAAAACCGCTCATCTATATAAGCTGTGTGATCATCG 780
QY 953 AGTGTGTGTTGAGGGTGAAGGCTCAGCAACCTGTATGTCTGCAACTACCAAGAAAGCT 1012
Db 781 AGTGTGTGTTGAGGGTGAAGGCTCAGCAACCTGTATGTCTGCAACTACCAAGAAAGCT 840
QY 1013 TCCCTGAGATGAGCTCATGACCTGTCAGTTTGTCTGTCCTGACACCTGCGGCGCAGAG 1072
Db 841 TCCCTGAGATGAGCTCATGACCTGTCAGTTTGTCTGTCCTGACACCTGCGGCGCAGAG 900
QY 1073 TCTCTTCTCTCAACTTCAACCTCTCCAACTGTGAGAGAAAGAGAGCGGGTTGAATACT 1132
Db 901 TCTCTTCTCTCAACTTCAACCTCTCCAACTGTGAGAGAAAGAGAGCGGGTTGAATACT 960
QY 1133 ACATCCCGGGCTCCACCAACCCCGAGGTTCAGAGCTGAGAGCAAGACCTTGGA 1192
Db 961 ACATCCCGGGCTCCACCAACCCCGAGGTTCAGAGCTGAGAGCAAGACCTTGGA 1020
QY 1193 ACATGCGGGGAACTTCAACCTCTCTCTGCAAGCTGTGACCAAGATGCCCAAGTCCAG 1252
Db 1021 ACATGCGGGGAACTTCAACCTCTCTCTGCAAGCTGTGACCAAGATGCCCAAGTCCAG 1080
QY 1253 GATTCCTCCGGCTGACAGTTTGTGTCCAACATCCCAAAATGAAAGCAATMAA 1312
Db 1081 GATTCCTCCGGCTGACAGTTTGTGTCCAACATCCCAAAATGAAAGCAATMAA 1140
QY 1313 TCTAGTGTGTTGACTTGAATGAGGAGCAGTGTCACTACCATGAGCCAGGCCG 1372
Db 1141 TCTAGTGTGTTGACTTGAATGAGGAGCAGTGTCACTACCATGAGCCAGGCCG 1200
QY 1373 TCAAAAGAGAGCCGCAAGTTTGTCCCTGAGCTGTGTCAGAAATCTCGGACTGCA 1432
Db 1201 TCAAAAGAGAGCCGCAAGTTTGTCCCTGAGCTGTGTCAGAAATCTCGGACTGCA 1260
QY 1433 GTAGCAACCTCAACCTGACATCTGGCTCCAAACAAATCTCCTTCTGTGATGATC 1492
Db 1261 GTAGCAACCTCAACCTGACATCTGGCTCCAAACAAATCTCCTTCTGTGATGATC 1320
QY 1493 TGAACGCTGTGATGATGTAATGTGAAAAAAACATAAGCTGACAGACACCGGTACTGCC 1552
Db 1321 TGAACGCTGTGATGATGTAATGTGAAAAAAACATAAGCTGACAGACACCGGTACTGCC 1380
QY 1553 AAAGAAATCTTACTCACTCAGGTGCGCAGTGTGACCTCTGACCTGCTGTGAGAGTGC 1612
Db 1381 AAAGAAATCTTACTCACTCAGGTGCGCAGTGTGACCTCTGACCTGCTGTGAGAGTGC 1440
QY 1613 ATGACTTCTCTGGAAGCTGTGATGCGCAAGGACAGGCTCAGCTGTGATGATGAG 1672
Db 1441 ATGACTTCTCTGGAAGCTGTGATGCGCAAGGACAGGCTCAGCTGTGATGATGAG 1500
QY 1673 CCGGAGAGCTGACAGCAGATACACAGAGAAAGCCTGCAACACAGCTTCAAGTACTCG 1732
Db 1501 CCGGAGAGCTGACAGCAGATACACAGAGAAAGCCTGCAACACAGCTTCAAGTACTCG 1560
QY 1733 TGGCCAGTGTGATACCCAGCAGAGCTGTGATCTTGGCTCTTCTGCGCGGAGGCTCTA 1792
Db 1561 TGGCCAGTGTGATACCCAGCAGAGCTGTGATCTTGGCTCTTCTGCGCGGAGGCTCTA 1620
QY 1793 TCAAGCAGATCGAGTGAAGCAGAACTCTCGGTGACCTTGGCAGCTTTGGCCCAAGCT 1852
Db 1621 TCAAGCAGATCGAGTGAAGCAGAACTCTCGGTGACCTTGGCAGCTTTGGCCCAAGCT 1680

QY 1853 TCAACAAAGAGGCTTCAGAGCAGGCTGTGACGCTCTCTTATACCTTATTTCAAGAG 1912
Db 1681 TCCGACAAAGAGGCTTCAGAGCAGGCTGTGACGCTCTCTTATACCTTATTTCAAGAG 1740
QY 1913 AAGGCGTTTTCAGGTTGACCCCTGACACAAAAGCAAGGCTACCTGAGGAGCCCAACT 1972
Db 1741 AAGGCGTTTTCAGGTTGACCCCTGACACAAAAGCAAGGCTACCTGAGGAGCCCAACT 1800
QY 1973 GGGACCGGGGCTGACATCCCTCACTCTGTGTCTGTGAAACATCAGCGTGCACAGACC 2032
Db 1801 GGGACCGGGGCTGACATCCCTCACTCTGTGTCTGTGAAACATCAGGTGTGCCAGAGACC 1860
QY 2033 AGTGGCTGCTGACTTCTTTAAGAGCGGAGCGGCTGTGCGCAGACAGGCGCG 2092
Db 1861 AGTGGCTGCTGACTTCTTTAAGAGCGGAGCGGCTGTGCGCAGACAGGCGCG 1920
QY 2093 CATTCATGATCATTCAGAGAGCGGAGCCCGGCTGAGAGATCTTACGCTGAGCAGAG 2152
Db 1921 CATTCATGATCATTCAGAGAGCGGAGCCCGGCTGAGAGATCTTACGCTGAGCAGAG 1980
QY 2153 ATGTGCTCCCAAGCCAGCTTCACATCAGAGCTTCTGGTCAACATCTCTAACTGCA 2212
Db 1981 ATGTGCTCCCAAGCCAGCTTCACATCAGAGCTTCTGGTCAACATCTCTAACTGCA 2040
QY 2213 GCCCAGAGCGGCAAGCAGTGAACCTGCTTTCTCGGTGACATTTACCCAGAGACTG 2272
Db 2041 GCCCAGAGCGGCAAGCAGTGAACCTGCTTTCTCGGTGACATTTACCCAGAGACTG 2100
QY 2273 TGGACTGATGCTCATCTCCATTCGAGGCGGTGGAGTGGAGTCTTACGCTGTGCGCC 2332
Db 2101 TGGACTGATGCTCATCTCCATTCGAGGCGGTGGAGTGGAGTCTTACGCTGTGCGCC 2160
QY 2333 TCGGGCTCATCTTGTCTGTGAAAAAGAAAGAAAGCAAAACAGGGCCCCGCTG 2392
Db 2161 TCGGGCTCATCTTGTCTGTGAAAAAGAAAGAAAGCAAAACAGGGCCCCGCTG 2220
QY 2393 TGGGATCTTACATGAGCAACATCAATCTGAGATGCCAGGACCCAAAAAGTTTCAGA 2452
Db 2221 TGGGATCTTACATGAGCAACATCAATCTGAGATGCCAGGACCCAAAAAGTTTCAGA 2280
QY 2453 AAGGGGAAAGGACAAATGCTCCATGCTATGAGTCACTGAGGACACCATGATATAG 2512
Db 2281 AAGGGGAAAGGACAAATGCTCCATGCTATGAGTCACTGAGGACACCATGATATAG 2340
QY 2513 GGCATCTGCTTACAGATTCAGAGGCTCTTCTGACGACAGAGTGTGACACTACCGG 2572
Db 2341 GGCATCTGCTTACAGATTCAGAGGCTCTTCTGACGACAGAGTGTGACACTACCGG 2400
QY 2573 CTTTCCAGGGCACATGGGGGTCTGTCTCTCCCTCCACCAACATATGCTCCAGGGCC 2632
Db 2401 CTTTCCAGGGCACATGGGGGTCTGTCTCTCCCTCCACCAACATATGCTCCAGGGCC 2460
QY 2633 CAACCTGCAAAAGTTGGCCATGAGAGCACTCTGCTGCTCCCTCTGTGATGAGAGTG 2692
Db 2461 CAACCTGCAAAAGTTGGCCATGAGAGCACTCTGCTGCTCCCTCTGTGATGAGAGTG 2520
QY 2693 AACCGTACACTTCTCCATCCCAAAATGGGATGTAAAGCAGAAAGCAGACAGACTTC 2752
Db 2521 AACCGTACACTTCTCCATCCCAAAATGGGATGTAAAGCAGAAAGCAGACAGACTTC 2580
QY 2753 CCTTACTGAACACTCAGAGGCCATGAGACACCGGCTGCTGCTTCACTTCCAGAGCCG 2812
Db 2581 CCTTACTGAACACTCAGAGGCCATGAGACACCGGCTGCTGCTTCACTTCCAGAGCCG 2640
QY 2813 TTTTGTGATTTTCAATTAAGAGAGGAGCTGAGACACCGGCTGCTGCTTCACTTCCAGAG 2872
Db 2641 TTTTGTGATTTTCAATTAAGAGAGGAGCTGAGACACCGGCTGCTGCTTCACTTCCAGAG 2700
QY 2873 CTAAAGAGAGGATTTATCAAGAGAAACAGCAGAGAGTTTCTGTGACACCGCCAACTT 2932
Db 2701 CTAAAGAGAGGATTTATCAAGAGAAACAGCAGAGAGTTTCTGTGACACCGCCAACTT 2760
QY 2933 CAATTTGCTCAGTGAATCATTTCTAAGGGCAAGACATGMAAATGATGAATTCATCTTG 2992

D	b	2761	CACATGTGCTAGTGGACTCTATTCTTAAGGCAAGACATTGAAATGATGAATTCGAAATTCG	2820
Q	y	2993	GATA CAGTCATGACAGCTCATGTGCTCCTCAACTTAAAGCTGTGGGTTAGCCAGCCTGTA	3052
D	b	2821	GATACAGTCATGACAGCTCATGTGCTCCTCAACTTAAAGCTGTGGGTTAGCCAGCCTGTA	2880
Q	y	3053	ATGAGAGAGAGAGAGGCGCTGAGTCACCTTAGCATAGGGTTGACGCAAGCCCTGGATTCAAG	3112
D	b	2881	ATGAGAGAGAGAGAGGCGCTGAGTCACCTTAGCATAGGGTTGACGCAAGCCCTGGATTCAAG	2940
Q	y	3113	TGTTAAACAGAGGCGCTGTGCGCTTCAGGACAAAGCTTCATTCGACAGGACCTACCTGA	3172
D	b	2941	TGTTAAACAGAGGCGCTGTGCGCTTCAGGACAAAGCTTCATTCGACAGGACCTACCTGA	3000
Q	y	3173	GGTCCCTTACTCTCACTGGGGTCCCAGAGTAAAAACGAATGTGCTTTTAAATTAATAT	3232
D	b	3001	GGTCCCTTACTCTCACTGGGGTCCCAGAGTAAAAACGAATGTGCTTTTAAATTAATAT	3060
Q	y	3233	TTATTTGGTGGTCCGTGTCTTATTTTAAAGATCAAAATGATTAACAACCTAGCTCTTTTAC	3292
D	b	3061	TTATTTGGTGGTCCGTGTCTTATTTTAAAGATCAAAATGATTAACAACCTAGCACTTTTAC	3120
Q	y	3293	CTGACTTAGTAATTAATCACTACTAATCGTGTGGAATGCTCGGGTTGACCTTCACTGAC	3352
D	b	3121	CTGACTTAGTAATTAATCACTACTAATCGTGTGGAATGCTCGGGTTGACCTTCACTGAC	3180
Q	y	3353	CGCTAGATTAACCGTGTGCTGTGCCCGCAGTGTGGGAATTAATTTACAACTGTGTCAAC	3412
D	b	3181	CGCTAGATTAACCGTGTGCTGTGCCCGCAGTGTGGGAATTAATTTACAACTGTGTCAAC	3240
Q	y	3413	AGAAAGAAATGTGTGTGTTTGAAGCAGCATGACACATATCTGCTTTGATTAAGACATTCC	3472
D	b	3241	AGAAAGAAATGTGTGTGTTTGAAGCAGCATGACACATATCTGCTTTGATTAAGACATTCC	3300
Q	y	3473	TGATTTCTAGTGGGTTCCGTGTTATCCCATGTGTGAAATTCATCTTGATATCCCATTTGT	3532
D	b	3301	TGATTTCTAGTGGGTTCCGTGTTATCCCATGTGTGAAATTCATCTTGGAATCCCATTTGT	3360
Q	y	3533	CCTATAGCTCAGCAATTAAGAAATTTCCCTCAAGTTTCCATGTGCGGTTCTCTAGCTG	3592
D	b	3361	CCTATAGCTCAGCAATTAAGAAATTTCCCTCAAGTTTCCATGTGCGGTTCTCTAGCTG	3420
Q	y	3593	CAGCAATATCTTGAACATTTAAAGAAATTTAGAAATATTCATCTCTTAAAAATGTT	3652
D	b	3421	CAGCAATATCTTGAACATTTAAAGAAATTTAGAAATATTCATCTCTCTTAAAAATGTT	3480
Q	y	3653	TAAATATATACCAACAGTGGCCCCCTGCAATTAGTTTCTGTGCACTGCAACCCATTTA	3712
D	b	3481	TAAATATATACCAACAGTGGCCCCCTGCAATTAGTTTCTGTGCACTGCAACCCATTTA	3540
Q	y	3713	CTTGATGACTTAAAAAACAACATTTAGCTTATAGTCCCGGGGATCAGAAATTCGAAAATGG	3772
D	b	3541	CTTGATGACTTAAAAAACAACATTTAGCTTATAGTCCCGGGGATCAGAAATTCGAAAATGG	3600
Q	y	3773	ATGTCCCTGATGAAATCAAGAGTGTGACGAGCTGTGCTCTTCTGAAGGCTCTAGGG	3832
D	b	3601	ATGTCCCTGATGAAATCAAGAGTGTGACGAGAGTGTGCTCTTCTGAAGGCTCTAGGG	3660
Q	y	3833	AGAAAGCGGTTCTTTGCCATTCAAAGCTTCTTAAAGGCTGCTGCATTTCCAGGCTCCAGT	3892
D	b	3661	AGAAAGCGGTTCTTTGCCATTCAAAGCTTCTTAAAGGCTGCTGCATTTCCAGGCTCCAGT	3720
Q	y	3893	GGCTGTGCAAGCTTTTCTCATGAGGATCATGTGACATGTGGGCTCCCATTCGCTCTT	3952
D	b	3721	GGCTGTGCAAGCTTTTCTCATGAGGATCATGTGACATGTGGGCTCCCATTCGCTCTT	3780
Q	y	3953	TGACTTACAAAGCCACCAAGAGATCCAGGATATCTTCTCATTAAGATCTCTTATC	4012
D	b	3781	TGACTTACAAAGCCACCAAGAGATCCAGGATATCTTCTCATTAAGATCTCTTATC	3840
Q	y	4013	ATCTCGAAGAGCTTTTGGCATGACAGACATATGCCAGGTGGGATTTAGACACAG	4072

[illegible]

```
QY 5152 GGTTCCTGCCAGTGAACCTTGCCCTTGTGAGCCGAGATGCTGACCTTCATTAAGAGGCC 5211
DB 4978 GGTTCCTGCCAGTGAACCTTGCCCTTGTGAGCCGAGATGCTGACCTTCATTAAGAGGCC 5037
QY 5212 AAGGAGGGGTGGGGCTTCCCTCCACAGAGAGCCCTTATTTGAATTCATCTGTGTG 5271
DB 5038 AAGGAGGGGTGGGGCTTCCCTCCACAGAGAGCCCTTATTTGAATTCATCTGTGTG 5097
QY 5272 AGCCCTAGCCCTCATCTTCAGATTCGCCCAACCTCCAGCCCTTCAAGCAGACTAG 5331
DB 5098 AGCCCTAGCCCTCATCTTCAGATTCGCCCAACCTCCAGCCCTTCAAGCAGACTAG 5157
QY 5332 GTGCTCTGCATTCACCCCAAGGTGGATTTGGCTTCTTAAGCTGTGCTACTGTCAACAT 5391
DB 5158 GTGCTCTGCATTCACCCCAAGGTGGATTTGGCTTCTTAAGCTGTGCTACTGTCAACAT 5217
QY 5392 CACCGACATCAGTGTGGCCCTGCAAGGACACCAAGTGGCCATTTCTCTCAACAGAGAGCT 5451
DB 5218 CACCGACATCAGTGTGGCCCTGCAAGGACACCAAGTGGCCATTTCTCAACAGAGAGCT 5277
QY 5452 CAAATCTCTGACAAAGTTGCTGCTCTGAGACAGATATTTCTGAGCTGTGCTCAG 5511
DB 5278 CAAATCTCTGACAAAGTTGCTGCTCTGAGACAGATATTTCTGAGCTGTGCTCAG 5337
QY 5512 TGAAGGGGCCAGCTGAGGAACTTGCTCTTTCTTTAAAGCCCAAGGCCCACTTAA 5571
DB 5338 TGAAGGGGCCAGCTGAGGAACTTGCTCTTTCTTTAAAGCCCAAGGCCCACTTAA 5397
QY 5572 TAAATATTTACAGGCTCAGTGAAGAGTGAAGTGCATTTGTGAAGCTACTGATAG 5631
DB 5398 TAAATATTTACAGGCTCAGTGAAGAGTGAAGTGCATTTGTGAAGCTACTGATAG 5457
QY 5632 CAGCCCACTGCTCATCCAGTGTCTGCCATGCTCAAGAGAGGCGCAGCGCTGAGGA 5691
DB 5458 CAGCCCACTGCTCATCCAGTGTCTGCCATGCTCAAGAGAGGCGCAGCGCTGAGGA 5517
QY 5692 CTGCTCTCTAATGCTGTGCTCATTTGCAAGAGGAAAGGTCTCAAGAGAGTCAACTG 5751
DB 5518 CTGCTCTCTAATGCTGTGCTCATTTGCAAGAGGAAAGGTCTCAAGAGAGTCAACTG 5577
QY 5752 GGAAGAAGCAAGAGCCCAAGGATGAGCCCTGTGAAGGTTAGCAACTGTGTGTGTG 5811
DB 5578 GGAAGAAGCAAGAGCCCAAGGATGAGCCCTGTGAAGGTTAGCAACTGTGTGTGTG 5637
QY 5812 ATCTGAGTGTCTCAGTGAATATTTATTTCACTGAGATACCTTTTATGAGTGTG 5871
DB 5638 ATCTGAGTGTCTCAGTGAATATTTATTTCACTGAGATACCTTTTATGAGTGTG 5697
QY 5872 ATTCATTTCTGCTTTTAAATTAATGAATGACCAAAAAAAGATATCAAGCTGTTTA 5931
DB 5698 ATTCATTTCTGCTTTTAAATTAATGAATGACCAAAAAAAGATATCAAGCTGTTTA 5757
QY 5932 GTGCTCTGCTAATTTGCCCCCTGTGCTCAGTGAAGGCCCCGCTTCCAGTGTGACTGT 5991
DB 5758 GTGCTCTGCTAATTTGCCCCCTGTGCTCAGTGAAGGCCCCGCTTCCAGTGTGACTGT 5817
QY 5992 GACAGGCTCAGATGAGCTCAGAGATGCTCTTAATTTGAGTGAATACAGAAAGCA 6051
DB 5818 GACAGGCTCAGATGAGCTCAGAGATGCTCTTAATTTGAGTGAATACAGAAAGCA 5877
QY 6052 GAGCTTTGGATACAGATTTCTTCTCTTCAATTTGATGCGGTGACGTGTGAAGCAGATG 6111
DB 5878 GAGCTTTGGATACAGATTTCTTCTCTCTCTCAATTTGATGCGGTGACGTGTGAAGCAGATG 5937
QY 6112 TTTTGTCCGGAATAAATAATAGTCTTGAAGTCTCGCC 6152
DB 5938 TTTTGTCCGGAATAAATAATAGTCTTGAAGTCTCGCC 5978
```

```
RESULT 5
BD160235
LOCUS BD160235 5573 bp DNA linear PAT 17-JAN-2003
```

```
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD160235
VERSION BD160235.1 GI:27865993
KEYWORDS JP 2002191363-A/15078.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 5573)
AUTHORS Ota,T., Isegai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Isht,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002191363-A 15078 09-JUL-2002;
COMMENT HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002191363-A/15078
PD 09-JUL-2002
PE 28-JUL-2000 JP 2000280990
PI TOSHIO OTA,TAKAO ISEGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KOJI
PI SAITO,K.
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
PI KEIICHI NAGAI,TETSUJI OTSUKI
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10,
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N5/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH key
Location/Qualifiers (275)..(2221).
FEATURES
source 1..5573
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 1373 a 1524 c 1290 g 1386 t
ORIGIN
Query Match 77.7%; Score 4786; DB 6; Length 5573;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 5566; Conservative 0; Mismatches 6; Indels 4; Gaps 4;
QY 569 AATATGACGTATATGTCAGAGCCCATGCTCTTTGGGAGAGTTGAGCTTCAGCCCTGACAT 628
DB 1 AATATGACGTATATGTCAGAGCCCATGCTCTTTGGGAGAGTTGAGCTTCAGCCCTGACAT 60
QY 629 CATTGTGCTACCTCTCAACAGAACTTTCATCTGAGATGTCAAACTCATTAAGACATG 688
DB 61 CATTGTGCTACCTCTCAACAGAACTTTCATCTGAGATGTCAAACTCATTAAGACATG 120
QY 689 GTTTAGAGCTGAGTTTTCATCTCTGCTGAGGCGAGATGCTCGGGTGAAGCTGCC 748
DB 121 G-TTAGAGCTGAGTTTTCATCTCTGCTGAGGCGAGATGCTCGGGTGAAGCTGCC 179
QY 749 CAGAGGAGTCACTACCTCCATCAGAGGCGCAATGAGTGAAGCCAGCTGTGACAGATCGAA 808
DB 180 CAGAGGAGTCACTACCTCCATCAGAGGCGCAATGAGTGAAGCCAGCTGTGACAGATCGAA 239
QY 809 CTTTGTGAGCAATGAGCTGTGCTCCGATCAAGATCAAGATCAAGAGAGTGAATAGGCT 868
DB 240 CTTTGTGAGCAATGAGCTGTGCTCCGATCAAGATCAAGATCAAGAGAGTGAATAGGCT 299
QY 869 TACACCTCCATGATTTCAAGCCCAAGAAATGTCCTCGGCTTTCAGATTTGCAACCGCTCAT 928
DB 300 TACACCTCCATGATTTCAAGCCCAAGAAATGTCCTCGGCTTTCAGATTTGCAACCGCTCAT 359
QY 929 CTAATAAACGTGTGTCATCATGAGTCTGTGTTGAGGAGTGAAGAGTCAAGACCTCTGA 988
DB 360 CTAATAAACGTGTGTCATCATGAGTCTGTGTTGAGGAGTGAAGAGTCAAGACCTCTGA 419
QY 989 TGTCTGCCAATCAAGAGGCTTCCCTGAGAGATGAGCTCATGAGTGTGAGTGTGTG 1048
DB 420 TGTCTGCCAATCAAGAGGCTTCCCTGAGAGATGAGCTCATGAGTGTGAGTGTGTG 479
```


1049 TTCTGCAACACTGCGGCGCAAGCTCTCTTCTCACTTCAACTCTTCAACTGTGAGA 1108
1109 GGAAGGAGGAGGGGTGAATCTACTATCCGGGGCTCCACCAACCCCGAGGTGTTCA 1168
540 GGAAGGAGGAGGGGTGAATCTACTATCCGGGGCTCCACCAACCCCGAGGTGTTCA 599
1169 AGCTGGAGGACAAGCAGCTGGGACATGCGGGGAATTTCAACTCTCTCTGCAAGGCT 1228
600 AGCTGGAGGACAAGCAGCTGGGACATGCGGGGAATTTCAACTCTCTCTGCAAGGCT 659
1229 GTGACCAAGATGCCCCAAGTCCAGGGATCTCCGGCTGCAATTTCAAGTTTGTGTCAC 1288
660 GTGACCAAGATGCCCCAAGTCCAGGGATCTCCGGCTGCAATTTGTGTCAC 719
1289 ATCCCAAAATGAAGCAATTAATTTAGTGTGTTAATTGAGTAAATGAGCAAGCATGT 1348
720 ATCCCAAAATGAAGCAATTAATTTAGTGTGTTAATTGAGTAAATGAGCAAGCATGT 779
1349 CACTGCAATGAGCGGCGGCGTCAAAACAGAGCGCAAGTTTGTCCCTGGCTGTTTGG 1408
780 CACTGCAATGAGCGGCGGCGTCAAAACAGAGCGCAAGTTTGTCCCTGGCTGTTTGG 839
1409 TGTGTCTAGAATCTGCGACCTGACGTAGCAACTCACTGACATCTGGCTCCAAACACA 1468
840 TGTGTCTAGAATCTGCGACCTGACGTAGCAACTCACTGACATCTGGCTCCAAACACA 899
1469 AAATCTCTTCTCTTGTGTGATCTGACACGTCTGTGTGATGAAATGTGAAAAAATCATTA 1528
900 AAATCTCTTCTCTTGTGTGATCTGACACGTCTGTGTGATGAAATGTGAAAAAATCATTA 959
1529 GGTGCAACAGACCAAGGCTGCTGCAAGAAATCTCACTCACTCCAGGTCAGTGAACA 1588
960 GGTGCAACAGACCAAGGCTGCTGCAAGAAATCTCACTCACTCCAGGTCAGTGAACA 1019
1589 TCTCTCACTGCTGTGTGAGCTGTGACATCTTCTCTGAAAGCTGTGTGTCCTCAAGACA 1648
1020 TCTCTCACTGCTGTGTGAGCTGTGACATCTTCTCTGAAAGCTGTGTGTCCTCAAGACA 1079
1649 GGTCTCACTGCTGTGTGAGCTGTGACATCTTCTCTGAAAGCTGTGTGTCCTCAAGACA 1708
1080 GGTCTCACTGCTGTGTGAGCTGTGACATCTTCTCTGAAAGCTGTGTGTCCTCAAGACA 1139
1709 GGAACACAGAGCTTCACTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1768
1140 GGAACACAGAGCTTCACTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1199
1769 GGTCTCTTCTGCGGAGGCTCTATCAAGCAGATCCAGGTGAAGCAGAACATCTCGGTGA 1828
1200 GGTCTCTTCTGCGGAGGCTCTATCAAGCAGATCCAGGTGAAGCAGAACATCTCGGTGA 1259
1829 CCCTTTCGCACTTTTGGCCCCAGCTTCCAAACAAGGCTTCCAGGCGAGGCTGTGACGCTGT 1888
1260 CCCTTTCGCACTTTTGGCCCCAGCTTCCAAACAAGGCTTCCAGGCGAGGCTGTGACGCTGT 1319
1889 CCTTTATACCTTATTTCAAGAAGGAGGCTTTTCAACGCTGACCCCTGACAAAAAGCA 1948
1320 CCTTTATACCTTATTTCAAGAAGGAGGCTTTTCAACGCTGACCCCTGACAAAAAGCA 1379
1949 AGGTCTACCTGAGGACCCCAACTGGGACCGGGGCTGTGCATCCCTCACTCTGTGTCTCT 2008
1380 AGGTCTACCTGAGGACCCCAACTGGGACCGGGGCTGTGCATCCCTCACTCTGTGTCTCT 1439
2009 GGAACATCAGCGTGTCCAGAGACCAAGTGGCTGTGCTGCTGCTTCTTTTAAAGACGAGCG 2068
1440 GGAACATCAGCGTGTCCAGAGACCAAGTGGCTGTGCTGCTGCTTCTTTTAAAGACGAGCG 1499
2069 GGTGTGTCTGCGACAGAGGCGCGCATTTATGATATCTCAAGACAGCGGACCCGGGCTGT 2128
1500 GGTGTGTCTGCGACAGAGGCGCGCATTTATGATATCTCAAGAGGACCGGACCCGGGCTGT 1559
2129 AGGAGATCTTCAAGCTGTGACGAGGATGTCTCCCAAGCAAGCTTCCACATCTCAGCT 2188

1560 AGGAGATCTTCAAGCTGTGACGAGGATGTCTCCCAAGCAAGCTTCCACATCTCAGCT 1619
2189 TCTGGGTCAACATCTCTTAATCTGACAGCCCAAGAGCGGCAAGAGCTAGACCTCTCTCT 2248
1620 TCTGGGTCAACATCTCTTAATCTGACAGCCCAAGAGCGGCAAGAGCTAGACCTCTCTCT 1679
2249 CGGTGACACTTACCCCAAGGACTGTGACCTTGAATCTGATCTCTCACTGCGAGGCTGTGAG 2308
1680 CGGTGACACTTACCCCAAGGACTGTGACCTTGAATCTGATCTCTCACTGCGAGGCTGTGAG 1739
2309 GTGAGTCTTACTGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2368
1740 GTGAGTCTTACTGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1799
2369 AGAAGACAAAGAGGCGCGCTGTGGGATATCAAGTGGCAACATCAATCTGAGATGC 2428
1800 AGAAGACAAAGAGGCGCGCTGTGGGATATCAAGTGGCAACATCAATCTGAGATGC 1859
2429 CGAGGACGCAAAAAAGTTTCAAGAAAGGCGAAGGACATGACTCCCATGTGTATGACAG 2488
1860 CGAGGACGCAAAAAAGTTTCAAGAAAGGCGAAGGACATGACTCCCATGTGTATGACAG 1919
2489 TCATGAGGACACATGTGTATATGCGCATCTGCTCAAGATTTCCAGCGCTCTCTCTGTC 2548
1920 TCATGAGGACACATGTGTATATGCGCATCTGCTCAAGATTTCCAGCGCTCTCTCTGTC 1979
2549 AGCCAGAGTGTGACACATACCGGCGCTTCCAGGGACCATGAGGAGCTGTCTCTCTCTCC 2608
1980 AGCCAGAGTGTGACACATACCGGCGCTTCCAGGGACCATGAGGAGCTGTCTCTCTCTCC 2039
2609 CACCCACATATGCTTCCAGGGCGCCCACTGCAAGTTGGGCACTGAGGAGCCACTCTCTC 2668
2040 CACCCACATATGCTTCCAGGGCGCCCACTGCAAGTTGGGCACTGAGGAGCCACTCTCTC 2099
2669 GGTCTCTCTCTGAGTCTGAGATGAAACGCTACACCTTCTCCATCCAAATGAGGAGATG 2728
2100 GGTCTCTCTCTGAGTCTGAGATGAAACGCTACACCTTCTCCATCCAAATGAGGAGATG 2159
2729 TAAAGACAGGACACACATCTTCCCTTACTGAAACCTCAGAGGCCATGAGGCCAGAG 2788
2160 TAAAGACAGGACACACATCTTCCCTTACTGAAACCTCAGAGGCCATGAGGCCAGAG 2219
2789 AATACTGATCATCTTCCAGAGCTTGTGTGATTTCAATTAAGAGGAGCACTGAGACACC 2848
2220 AATACTGATCATCTTCCAGAGCTTGTGTGATTTCAATTAAGAGGAGCACTGAGACACC 2279
2849 CGTCCGTGTTCTTAAACAGAAATCTTAAAGAGGAAATTAACAGAGGACACAGAGA 2908
2280 CGTCCGTGTTCTTAAACAGAAATCTTAAAGAGGAAATTAACAGAGGACACAGAGA 2339
2909 GGTCTTCTGAGACCGCCCACTTCACTGTCTGATGAGCTCACTTTAAGGCAAGACA 2968
2340 GGTCTTCTGAGACCGCCCACTTCACTGTCTGATGAGCTCACTTTAAGGCAAGACA 2399
2969 TTGAATAATGAAATTTCCAAATCTGAGTACAGTCAAGACGTCATGTGTCTCTCACTTA 3028
2400 TTGAATAATGAAATTTCCAAATCTGAGTACAGTCAAGACGTCATGTGTCTCTCACTTA 2459
3029 GGTCTGCGTTTAAAGCAGCTGTAAATGAGAGGAGGAGGCTGAGTCACTTGAAGTAAAGG 3088
2460 GGTCTGCGTTTAAAGCAGCTGTAAATGAGAGGAGGAGGCTGAGTCACTTGAAGTAAAGG 2519
3089 TTGACAGCAAGCTGTGATTTCAAGAGGCTTGGCTCTTCAAGCAACAGT 3148
2520 TTGACAGCAAGCTGTGATTTCAAGAGGCTTGGCTCTTCAAGCAACAGT 2579
3149 TTCAATTTCCAAAGAGCTTACCTGAGTCTCTCACTGTGGGTCTCCCAAGATGAAGAAC 3208
2580 TTCAATTTCCAAAGAGCTTACCTGAGTCTCTCACTGTGGGTCTCCCAAGATGAAGAAC 2639
3209 GACAAATGTCCTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 3268

D	b	2640	GACAAATGCGCTTTTATATATATATTTGTTGGTGGCTCGTGTTATTTTAAAGATCAAAAT	2639
Q	y	3369	GTATTAACCACTAGACTCTTTTCACTGACTTAGTATATACCTAATCACTAAGCTTTGGAT	3328
D	b	2700	GTATTAACCACTAGACTCTTTTCACTGACTTAGTATATACCTAATCACTAAGCTTTGGAT	2759
Q	y	3329	GCCGTGGGTGGTGAAGCTTTACTGACCGCTAAGTAAACGTGACCTGCTCCCAAGTGTGG	3388
D	b	2760	GCCGTGGGTGGTGAAGCTTTACTGACCGCTAAGTAAACGTGACCTGCTCCCAAGTGTGG	2819
Q	y	3389	GAATTAATTTACAATCTGTCCAAACAGAAAATAATGTGTGTTTGAGCAGATTGACACA	3448
D	b	2820	GAATTAATTTACAATCTGTCCAAACAGAAAATAATGTGTGTTTGAGCAGATTGACACA	2879
Q	y	3449	TATCTGCTTTGATTAAGACACTTCTGATTTCTAGTGGTTCGTGTATCCCATTTGTG	3508
D	b	2880	TATCTGCTTTGATTAAGACACTTCTGATTTCTAGTGGTTCGTGTATCCCATTTGTG	2939
Q	y	3509	GAATTTCACTTTGAAATCCCATTTGTCTTACTCTAGCAATTAAGAAATTTCTCAAGT	3568
D	b	2940	GAATTTCACTTTGAAATCCCATTTGTCTTACTCTAGCAATTAAGAAATTTCTCTCAAGT	2999
Q	y	3569	TTCCATGATGGGGTTCCTCTAGCTGAGCAATTAATTGACATTTTAAAGAAATTTAGAGA	3628
D	b	3000	TTCCATGATGGGGTTCCTCTAGCTGAGCAATTAATTGACATTTTAAAGAAATTTAGAGA	3059
Q	y	3629	ATATTTCTCATCTCTAATAAATGTTTAAATATATACAAACAGTGGCCCTCGCATTAAGTT	3688
D	b	3060	ATATTTCTCATCTCTAATAAATGTTTAAATATATACAAACAGTGGCCCTCGCATTAAGTT	3119
Q	y	3689	TTCTGTGTCGCACTGCAACCCATTACTTGTGTGCTTAAAAACAACAATTAGCTTATAGTC	3748
D	b	3120	TTCTGTGTCGCACTGCAACCCATTACTTGTGTGCTTAAAAACAACAATTAGCTTATAGTC	3179
Q	y	3749	CTGGGGGATCAAAATTCGAAAATGATGTCGCTGAAATGAAATTCAGGTGTAGCAGAGCT	3808
D	b	3180	CTGGGGGATCAAAATTCGAAAATGATGTCGCTGAAATGAAATTCAGGTGTAGCAGAGCT	3239
Q	y	3809	GTGCTCTCTTTCTGAAGGCTCTAGGGAGAACCGGTTCTTGGCAATTTCAAGCTTTCTAGAGG	3868
D	b	3240	GTGCTCTCTTTCTGAAGGCTCTAGGGAGAACCGGTTCTTGGCAATTTCAAGCTTTCTAGAGG	3299
Q	y	3869	CTGGCTGCATTTCCAGAGCTCCAGTGGCTGATGCTGATCAAGCTTTTCTCACATGGCATCACTGTGA	3928
D	b	3300	CTGGCTGCATTTCCAGAGCTCCAGTGGCTGATGCTGATCAAGCTTTTCTCACATGGCATCACTGTGA	3359
Q	y	3929	CACATGGCCCTCCCACTTCCTGCTTTGATTTACAAAGCCACACAGGAAATCCAGAGTATAT	3988
D	b	3360	CACATGGCCCTCCCACTTCCTGCTTTGATTTACAAAGCCACACAGGAAATCCAGAGTATAT	3419
Q	y	3989	CTCTCCATCTTAAAGATCCTTCATCATCTCTGGAGAAGCTTTTGGCATGCAAGACAATA	4048
D	b	3420	CTCTCCATCTTAAAGATCCTTCATCATCTCTGGAGAAGCTTTTGGCATGCAAGACAATA	3479
Q	y	4049	GCCACAGGTGGGGATTAGAGACCAAGACACTTTTGGGGTGTGTTATTTCTGCTTACCAAC	4108
D	b	3480	GCCACAGGTGGGGATTAGAGACCAAGACACTTTTGGGGTGTGTTATTTCTGCTTACCAAC	3539
Q	y	4109	CTTCTCTGCACTGACTCCCAAGAGAGGCTCTAATAAATGATCTGGCGACAGGATGTTT	4168
D	b	3540	CTTCTCTGCACTGACTCCCAAGAGAGGCTCTAATAAATGATCTGGCGACAGGATGTTT	3599
Q	y	4169	TGTTTAGCTTGGGACTCTTAACACTT-AAAAAACCACAGATCGAAGATCTGGCGATGC	4227
D	b	3600	TGTTTAGCTTGGGACTCTTAACACTTAAAAAACCACAGATCTGGCGATGCTGGCGATGC	3659
Q	y	4228	TGGGGCTTCAATTTCTCACTTAGCAACAATGGCTGGAGCTGGGACCAAGCTCTGCTTTTA	4287
D	b	3660	TGGGGCTTCAATTTCTCACTTAGCAACAATGGCTGGAGCTGGGACCAAGCTCTGCTTTTA	3719
Q	y	4288	GAAGGGGTGTCACTTCAACGAGGTCAACAAGCCCACTACGCGCTTATCACTTCCACA	4347
D	b	3720	GAAGGGGTGTCACTTCAACGAGGTCAACAAGCCCACTACGCGCTTATCACTTCCACA	3779

OY	4348	ATGAGCGTAAGTGTGTTGTTCTTCTACAGTCAATGCCCTGAGGTGATTTATGTGATG	4407
Db	3780	ATGAGCGTAAGTGTGTTGTTCTTCTACAGTCAATGCCCTGAGGTGATTTATGTGATG	3839
OY	4408	AAAAAGAAAGCTGSGATTAATCTCTATATAGGTAAGTAACATAGACCAATGTGTGC	4467
Db	3840	AAAAAGAAAGCTGSGATTAATCTCTATATAGGTAAGTAACATAGACCAATGTGTGC	3899
OY	4468	TCACATTACCCCTTTTCCTTTTTTCTTTTTTCTTTTTTCTTTTTTAAATGTGAGACA	4527
Db	3900	TCACATTACCCCTTTTCCTTTTTTCTTTTTTCTTTTTTCTTTTTTAAATGTGAGACA	3959
OY	4528	GGATCTATCTGTGTGCTTAGCGTGAAGTGACAGTGGGCAATCTCGGCTCACTGCAACT	4587
Db	3960	GGATCTATCTGTGTGCTTAGCGTGAAGTGACAGTGGGCAATCTCGGCTCACTGCAACT	4019
OY	4588	CTGCGCTCTGGGGCTCAAGCAATCTCCACTGACCTCCCAATTACTGGGATCACTGG	4647
Db	4020	CTGCGCTCTGGGGCTCAAGCAATCTCCACTGACCTCCCAATTACTGGGATCACTGG	4079
OY	4648	CACAAACACACATGCCAGCTAATTTTGTATTTTTGTAGAGACAGGGTTTCACATGTT	4707
Db	4080	CACAAACACACATGCCAGCTAATTTTGTATTTTTGTAGAGACAGGGTTTCACATGTT	4138
OY	4708	GCCCAAGCTGTCTCAACCTCTCTGGGCTCAAGCAATCTCTCTGTGGCTCCCAAGT	4767
Db	4139	GCCCAAGCTGTCTCAACCTCTCTGGGCTCAAGCAATCTCTCTGTGGCTCCCAAGT	4198
OY	4768	GCTGGGATTAAGATGTGAGCACCGGATCAAGGCCACACCCCATTTATACCAATTAC	4827
Db	4199	GCTGGGATTAAGATGTGAGCACCGGATCAAGGCCACACCCCATTTATACCAATTAC	4258
OY	4828	CTGCGCCAGTAAGTGTGAGACTTTTGCTTCTCTACCCCTGCTGTGATCTGGAAGAGAGGA	4887
Db	4259	CTGCGCCAGTAAGTGTGAGACTTTTGCTTCTCTCTACCCCTGCTGTGATCTGGAAGAGAGGA	4318
OY	4888	TTATGTATAGCTTGTGAGACAGTCCCAAGTTCAATTTTCTGGGGCAAAAATCTTCTT	4947
Db	4319	TTATGTATAGCTTGTGAGACAGTCCCAAGTTCAATTTTCTGGGGCAAAAATCTTCTT	4378
OY	4948	CAAAAATTAATATGACTTCTTATGTATTCATGAAATTCACTTGGAATTGCAACGGCTTCAA	5007
Db	4379	CAAAAATTAATATGACTTCTTATGTATTCATGAAATTCACTTGGAATTGCAACGGCTTCAA	4438
OY	5008	CTTGTTCACATGGGCTAATGAAAGAAATTTTATGCTCCTAATAGGCGTACTGCA	5067
Db	4439	CTTGTTCACATGGGCTAATGAAAGAAATTTTATGCTCCTAATAGGCGTACTGCA	4498
OY	5068	GACCTCTTGAACACTTTCAGAGAGATAGATATTTAAGTCATGCGCTTGGCGTTGCTAT	5127
Db	4499	GACCTCTTGAACACTTTCAGAGAGATAGATATTTAAGTCATGCGCTTGGCGTTGCTAT	4558
OY	5128	GGCACCTTTTCCCTTCTGAAAAGTCTGGTTCCTGCCAAGTACCTTGGGCTTTGTAGCCGA	5187
Db	4559	GGCACCTTTTCCCTTCTGAAAAGTCTGGTTCCTGCCAAGTACCTTGGGCTTTGTAGCCGA	4618
OY	5188	GATGCTGACCTGTGATTAAGGGGCAAAAGAGGGGTGGGCTTCTTCCCTCACTGAAGAG	5247
Db	4619	GATGCTGACCTGTGATTAAGGGGCAAAAGAGGGGTGGGCTTCTTCCCTCACTGAAGAG	4678
OY	5248	CCCTTATTTGAATTCATCTGTGTGGAGGCTTAGGCGCTCATTTCTGCAATTTCCCAACCTC	5307
Db	4679	CCCTTATTTGAATTCATCTGTGTGGAGGCTTAGGCGCTCATTTCTGCAATTTCCCAACCTC	4738
OY	5308	CCAGCCCCCTTTCAGAGCAGACTAGGTGCTCTGCAATTCACCCAAAGTGGGATTTGGCTTC	5367
Db	4739	CCAGCCCCCTTTCAGAGCAGACTAGGTGCTCTGCAATTCACCCAAAGTGGGATTTGGCTTC	4797
OY	5368	CTTAGGCTGGGCTACTTGTACATCAACCAATCATCTGTGTGCTGCAAGGACACACAGTG	5427
Db	4798	CTTAGGCTGGGCTACTTGTACATCAACCAATCATCTGTGTGCTGCAAGGACACACAGTG	4857

QY 1049 TTCTGACACCTGCGGGCCAGCGTCTCTTCTCACTTCAACTCTTCAACTCTGAGA 1108
Db 480 TTCTGACACCTGCGGGCCAGCGTCTCTTCTCACTTCAACTCTTCAACTCTGAGA 539
QY 1109 GGAAGAGAGAGGGGTTGAATACTACATCCGGGGCTCCACGACCAACCCCGAGGTGTA 1168
Db 540 GGAAGAGAGAGGGGTTGAATACTACATCCGGGGCTCCACGACCAACCCCGAGGTGTA 599
QY 1169 AGCTGGAGACAGACAGCTGGGACATGGCGGGAACTTCAACTCTCTGTCAGAGGT 1228
Db 600 AGCTGGAGACAGACAGCTGGGACATGGCGGGAACTTCAACTCTCTGTCAGAGGT 659
QY 1229 GTGACCAAGATGCCCAAGATCCAGGGATCTCGGCTGCACTTCCAGTTTGTGTCAC 1288
Db 660 GTGACCAAGATGCCCAAGATCCAGGGATCTCGGCTGCACTTCCAGTTTGTGTCAC 719
QY 1289 ATCCACAAATGAACATTAATTAATAGTGGTTGATTTAGTATGAGCGGACATGT 1348
Db 720 ATCCACAAATGAACATTAATTAATAGTGGTTGATTTAGTATGAGCGGACATGT 779
QY 1349 CACTCACATCGACGACCGGCGCCGTCAACAGAGCGCAAGTTTGTCCCTGGCTGTTTGC 1408
Db 780 CACTCACATCGACGACCGGCGCCGTCAACAGAGCGCAAGTTTGTCCCTGGCTGTTTGC 839
QY 1409 TGTGTATGAATCTGGACCTGACATGACAACTTCACTGACATCTGGCTCCAAACAA 1468
Db 840 TGTGTATGAATCTGGACCTGACATGACAACTTCACTGACATCTGGCTCCAAACAA 899
QY 1469 AATCTCTCTTCTTGTGATGATCTGACAGTGTGATGATGATGATGATGATGATGAT 1528
Db 900 AATCTCTCTTCTTGTGATGATCTGACAGTGTGATGATGATGATGATGATGATGAT 959
QY 1529 GCTGACACAGACACCGGTACTGCAAGAAATCTTACTACTCTGAGTGGCCAGTGA 1588
Db 960 GCTGACACAGACACCGGTACTGCAAGAAATCTTACTACTCTGAGTGGCCAGTGA 1019
QY 1589 TCTTCACTGCTGTGAGAGTGCATGACTTCTCTGAGAACTGTGTGTGTCAGAGAA 1648
Db 1020 TCTTCACTGCTGTGAGAGTGCATGACTTCTCTGAGAACTGTGTGTGTCAGAGAA 1079
QY 1649 GGTCAAGCTGTGTGAGAGTGCAGAGCTGAGAGCAATACAGAGAGAGAGAGAG 1708
Db 1080 GGTCAAGCTGTGTGAGAGTGCAGAGCTGAGAGCAATACAGAGAGAGAGAGAG 1139
QY 1709 GCAACACAGACTTCACTACTCTGTGAGAGTGCATACCAAGCAGAGACTGTACTTGC 1768
Db 1140 GCAACACAGACTTCACTACTCTGTGAGAGTGCATACCAAGCAGAGACTGTACTTGC 1199
QY 1769 GCTCTTCTGCGCGGAGAGCTTATCAAGAGATCCAGGTGAAGAGAACTCTGGTA 1828
Db 1200 GCTCTTCTGCGCGGAGAGCTTATCAAGAGATCCAGGTGAAGAGAACTCTGGTA 1259
QY 1829 CCGTTGGACCTTGGCGGAGTTCACAGAGGCTTCAGAGAGGCTGAGAGAGTGT 1888
Db 1260 CCGTTGGACCTTGGCGGAGTTCACAGAGGCTTCAGAGAGGCTGAGAGAGTGT 1319
QY 1889 CCTTTATACCTTATTTCAAGAGAGAGGCTTTCACAGGTCACCTGACCAAAAGCA 1948
Db 1320 CCTTTATACCTTATTTCAAGAGAGAGGCTTTCACAGGTCACCTGACCAAAAGCA 1379
QY 1949 AGGTCTACCTGAGAGACCCCAACTGGAGCCGGGGCTGACCTCTGCTGTGCTCT 2008
Db 1380 AGGTCTACCTGAGAGACCCCAACTGGAGCCGGGGCTGACCTCTGCTGTGCTCT 1439
QY 2009 GGAACATTCAGCGGCGGACAGAGTGGCTGAGAGCTTCTTTAAGAGACGGAGG 2068
Db 1440 GGAACATTCAGCGGCGGACAGAGTGGCTGAGAGCTTCTTTAAGAGACGGAGG 1499
QY 2069 GCGTGTCTGCGCAGACAGAGGCGGCACTTCAATGATTCAGAGACAGAGAGCCGGGCTG 2128
Db 1500 GCGTGTCTGCGCAGACAGAGGCGGCACTTCAATGATTCAGAGACAGAGAGCCGGGCTG 1559
QY 2129 AGGAGATCTTCAAGCTGAGAGAGATGTGCTCCCAAGCCAAAGCTTCAACATCAGAGCT 2188

Db 1560 AGGAGATCTTCAAGCTGAGAGAGATGTGCTCCCAAGCCAAAGCTTCAACATCAGAGCT 1619
QY 2189 TCTGGGTCAACATCTCTAAGTGCAGACCCCAAGAGCGGCAAGAGCTAGACTGTCTTCT 2248
Db 1620 TCTGGGTCAACATCTCTAAGTGCAGACCCCAAGAGCGGCAAGAGCTAGACTGTCTTCT 1679
QY 2249 CGGTGACACTTACCCCAAGAGCTGTGACCTGTATCTCTCATTCGACGGGTGGAG 2308
Db 1680 CGGTGACACTTACCCCAAGAGCTGTGACCTGTATCTCTCATTCGACGGGTGGAG 1739
QY 2309 GTGAGTCTTACTGCTGTGCTGCTCGGGCTCATATTTGCTGTGAGAAAGAGAA 2368
Db 1740 GTGAGTCTTACTGCTGTGCTGCTCGGGCTCATATTTGCTGTGAGAAAGAGAA 1799
QY 2369 AGAAGACAAACAGAGGAGCCGCTGTGGATCTAATGACATGACATCAATACTAGATG 2428
Db 1800 AGAAGACAAACAGAGGAGCCCGCTGTGGATCTAATGACATGACATCAATACTAGATG 1859
QY 2429 CGAGGACGCAAAAGTTTTCAGAAAGGCGAAGAGCAATGATCCCATGTGTATGAG 2488
Db 1860 CGAGGACGCAAAAGTTTTCAGAAAGGCGAAGAGCAATGATCCCATGTGTATGAG 1919
QY 2489 TCATGAGAGACCAATGTGATATGGGCAATGCTACAGAGATCCAGCGGCTCTCTG 2548
Db 1920 TCATGAGAGACCAATGTGATATGGGCAATGCTACAGAGATCCAGCGGCTCTCTG 1979
QY 2549 ACCAGAGGTGACACATACCGGCGTTCCAGAGGACATGAGGGGTCTGCTCCCTCC 2608
Db 1980 ACCAGAGGTGACACATACCGGCGTTCCAGAGGACATGAGGGGTCTGCTCCCTCC 2039
QY 2609 CACCAACATATGCTTCAGAGGCGCCCACTGCAAGTTGGCCACTGAGAGCCACTCTC 2668
Db 2040 CACCAACATATGCTTCAGAGGCGCCCACTGCAAGTTGGCCACTGAGAGCCACTCTC 2099
QY 2669 GCTCCCTCTGAGTCTGAGAGTGAACGTATACCTTCTCCATCCCAATGGGGATG 2728
Db 2100 GCTCCCTCTGAGTCTGAGAGTGAACGTATACCTTCTCCATCCCAATGGGGATG 2159
QY 2729 TAAAGACAGAGACACAGACTTCCCTTACTGAACCTGAGAGCCATGAGAGCCAGAG 2788
Db 2160 TAAAGACAGAGACACAGACTTCCCTTACTGAACCTGAGAGCCATGAGAGCCAGAG 2219
QY 2789 AATTAATGATCATTCACAGAGCTTGTGATGATTAAGAGAGGCACTGAGACAC 2848
Db 2220 AATTAATGATCATTCACAGAGCTTGTGATGATTAAGAGAGGCACTGAGACAC 2279
QY 2849 CGTCCGTGTTCTTAACAGAAATCTTAAGAGAGAGAAATTAACAGAGAGAGAGAG 2908
Db 2280 CGTCCGTGTTCTTAACAGAAATCTTAAGAGAGAGAAATTAACAGAGAGAGAGAG 2339
QY 2909 GGTTTTCTGACACCGCACTTCACTGCTCAGTGAATCTTCAAGGCAAGAGCA 2968
Db 2340 GGTTTTCTGACACCGCACTTCACTGCTCAGTGAATCTTCAAGGCAAGAGCA 2399
QY 2969 TTGAATAATGATTAATTCATGTGATACAGTCAATGATGTCTCTCAACTTA 3028
Db 2400 TTGAATAATGATTAATTCATGTGATACAGTCAATGATGTCTCTCAACTTA 2459
QY 3029 GGTGTGCGGTTAAGCAGCTGTATTAAGAGAGAGAGGCTGAGTCACTGAGATAGG 3088
Db 2460 GGTGTGCGGTTAAGCAGCTGTATTAAGAGAGAGAGGCTGAGTCACTGAGATAGG 2519
QY 3089 TTGACAGCAAGCCCTGATTCAGAGTGTAAACAGAGAGCTTGGCTTTCAGAGCAAGT 3148
Db 2520 TTGACAGCAAGCCCTGATTCAGAGTGTAAACAGAGAGCTTGGCTTTCAGAGCAAGT 2579
QY 3149 TCCAAATTCAGAGAGCTTCACTGAGTCCCTCATCTCATCTGAGGTCCCAAGATGA 3208
Db 2580 TCCAAATTCAGAGAGCTTCACTGAGTCCCTCATCTCATCTGAGGTCCCAAGATGA 2639
QY 3209 GACATGTGCTTTTATTTATTTATTTATTTGAGTCTGTGTATTTAAGAGATCAAT 3268

Dp	2640	AACAATGCGCTTTTATATTAATTATTTATTTGTTGGTGGTCTGTGTTATTTAAGAGATCAAT	2659
Qy	3289	GTATTAACCACTAGCTCTTTTCACTGACTTAAGTATTAATCTATATCTACTACTGTTGGAT	3328
Dp	2700	GTATTAACCACTAGCTCTTTTCACTGACTTAAGTATTAATCTATATCTACTACTGTTGGAT	2759
Qy	3329	GCCGGGGTGTGACTTCTACTGACCGCGTAGATAAAGCGTGCCGTGCCCCAGGCGTGCG	3388
Dp	2760	GCCGGGGTGTGACTTCTACTGACCGCGTAGATAAAGCGTGCCGTGCCCCAGGCGTGCG	2819
Qy	3389	GAATTAATTACAACTGTGTCCAACAGAAAATAATGTGTGTTTGGACAGCATTTGACA	3448
Dp	2820	GAATTAATTACAACTGTGTCCAACAGAAAATAATGTGTGTTTGGACAGCATTTGACA	2879
Qy	3449	TATCTGCTTTGATTAAGAGACTTCTGATTTCTTAAGTCGAGTTCGTGTTATCCATTGTG	3508
Dp	2880	TATCTGCTTTGATTAAGAGACTTCTGATTTCTTAAGTCGAGTTCGTGTTATCCATTGTG	2939
Qy	3509	GAATTAATCTTTGAATGCCATTGTCCTTAATCTTAAGTAAAGAAATTTCTCAAGT	3568
Dp	2940	GAATTAATCTTTGAATGCCATTGTCCTTAATCTTAAGTAAAGAAATTTCTCAAGT	2999
Qy	3569	TTCCATGTCGGGTCTCTCTAGCTGAGCAATCTTTGACATTTTAAAGAAATTTAGAG	3628
Dp	3000	TTCCATGTCGGGTCTCTCTAGCTGAGCAATCTTTGACATTTTAAAGAAATTTAGAG	3059
Qy	3629	ATATTTCTGATCCTCTAATAAATGTTTAAATATATACCAACAGTGGCCCCCTGCATTAAT	3688
Dp	3060	ATATTTCTGATCCTCTAATAAATGTTTAAATATATACCAACAGTGGCCCCCTGCATTAAT	3119
Qy	3689	TTCTGTTGGCACTGCACCCATTAATCTTGTAGCTTTAAAAACAACATTAAGCTTATAGTC	3748
Dp	3120	TTCTGTTGGCACTGCACCCATTAATCTTGTAGCTTTAAAAACAACATTAAGCTTATAGTC	3179
Qy	3749	CTGGGGATCAGAAATTCGAAAATGATGTCCCTGATGTAAGAAATCAAGGTCAGAGAGCT	3808
Dp	3180	CTGGGGATCAGAAATTCGAAAATGATGTCCCTGATGTAAGAAATCAAGGTCAGAGAGCT	3239
Qy	3809	GTCGTCCTTTCTGAAGGCTCTAAGGAGAAAGCCGGTTCCTTGCCATTTCAAGCTTTAAGG	3868
Dp	3240	GTCGTCCTTTCTGAAGGCTCTAAGGAGAAAGCCGGTTCCTTGCCATTTCAAGCTTTAAGG	3299
Qy	3869	CTGGCTGCAATTCACAGGCTCCAGTGGCTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCG	3928
Dp	3300	CTGGCTGCAATTCACAGGCTCCAGTGGCTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCG	3359
Qy	3929	CACGTGGCCCTCCCACTGCTGCTTTGAATTAAGGCCACAGAGAAATCCAGAGTAAAT	3988
Dp	3360	CACGTGGCCCTCCCACTGCTGCTTTGAATTAAGGCCACAGAGAAATCCAGAGTAAAT	3419
Qy	3989	CTTCTCATCTAAGATCTCTTATCATCTGTGAAGAGCTTTTGTGCTAGTCAAGACATTA	4048
Dp	3420	CTTCTCATCTAAGATCTCTTATCATCTGTGAAGAGCTTTTGTGCTAGTCAAGACATTA	3479
Qy	4049	GCACAGGTGGGATTAAGACCAAGACATCTTTGGGGTGTGTTATTTCTGCTACACAC	4108
Dp	3480	GCACAGGTGGGATTAAGACCAAGACATCTTTGGGGTGTGTTATTTCTGCTACACAC	3539
Qy	4109	CTTCTGCTGACTGACTCCCAAGAGAGGCTTCAAAAAGATCTTGGCCGACAGGATGTTT	4168
Dp	3540	CTTCTGCTGACTGACTCCCAAGAGAGGCTTCAAAAAGATCTTGGCCGACAGGATGTTT	3599
Qy	4169	TGTTTAGCTTGGGACTTAACACTT-AAAAAAACCCAGATCAAGAAATCTGGCCATGC	4227
Dp	3600	TGTTTAGCTTGGGACTTAACACTTAACTTAABAAAAAACCCAGATCAAGAAATCTGGCCATGC	3659
Qy	4228	TGGGGCTCACTTCTCACTAGCAACACTGGCTGGAGCTGGGACCAAGCTCTGCTTTA	4287
Dp	3660	TGGGGCTCACTTCTCACTAGCAACACTGGCTGGAGCTGGGACCAAGCTCTGCTTTA	3719
Qy	4288	GAAGGGGTGTCACTTCAACAGATTCACAAGCCCACTAGCGCCCTATATCACTTCCACA	4347
Dp	3720	GAAGGGGTGTCACTTCAACAGATTCACAAGCCCACTAGCGCCCTATATCACTTCCACA	3779

OY	4348	ATGAGGCTAAGTGTGTTGTTTCTAATGATCAATGCGCCGTGAGGGTGCATTTATGTATG	4407
OY	4348	ATGAGGCTAAGTGTGTTGTTTCTAATGATCAATGCGCCGTGAGGGTGCATTTATGTATG	4407
Db	3780	ATGAGGCTAAGTGTGTTGTTTCTAATGATCAATGCGCCGTGAGGGTGCATTTATGTATG	3839
OY	4408	AAAAAAGAAGCTGGATTTATCTCTATCATGAGTATAGACATAGACCAATGTGTGC	4467
Db	3840	AAAAAAGAAGCTGGATTTATCTCTATCATGAGTATAGACATAGACCAATGTGTGC	3899
OY	4468	TCACATTACCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTATGTAGACA	4527
Db	3900	TCACATTACCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTATGTAGACA	3959
OY	4528	GGATCTATCTGTGTGCTAGGCGTGAAGTCAAGTGGGGAATCTGGCTACGTGCACCT	4587
Db	3960	GGATCTATCTGTGTGCTAGGCGTGAAGTCAAGTGGGGAATCTGGCTACGTGCACCT	4019
OY	4588	CTGCCTCTGGGGCTCAAGCAATTTCCACCTCCAGCTCCCAATATAGCTGGATCATCTG	4647
Db	4020	CTGCCTCTGGGGCTCAAGCAATTTCCACCTCCAGCTCCCAATATAGCTGGATCATCTG	4079
OY	4648	CACAAACCAACATGCCAGCTAATTTTGTATTTTGTAGAGACAGGGTTTACCATGTT	4707
Db	4080	CACAAACCAACATGCCAGCTAATTTTGTATTTTGTAGAGACAGGGTTTACCATGTT	4138
OY	4708	GCCCAAGCTGTCTCAACCTCTGTGGGCTCAAGCAATCTCTCTGTGTGGCTCCCAAGT	4767
Db	4139	GCCCAAGCTGTGTCTCAACCTCTGTGGGCTCAAGCAATCTCTGTGTGGCTCCCAAGT	4198
OY	4768	GCTGGGATTTAGATGTAGAGCAACGATCCAGCCGACACCCCATTTTATACCAATTAC	4827
Db	4199	GCTGGGATTTAGATGTAGAGCAACGATCCAGCCGACACCCCATTTTATACCAATTAC	4258
OY	4828	CTGCCAGTAACTGTGACCTTTTGTCTTCAACCCCTGTCTGTATCTGAAAGAGAGGA	4887
Db	4259	CTGCCAGTAACTGTGACCTTTTGTCTTCAACCCCTGTCTGTATCTGAAAGAGAGGA	4318
OY	4888	TTATGTATTAGCTTGTGACGACAGTCCCAAGTTCAATATTTCTGGGCAAAACTTCTT	4947
Db	4319	TTATGTATTAGCTTGTGACGACAGTCCCAAGTTCAATATTTCTGGGCAAAACTTCTT	4378
OY	4948	CAAAAAATTAATGTACTTCAATGTATTCATGTGAATTCACCTTGGAAATGACACGGCTCAA	5007
Db	4379	CAAAAAATTAATGTACTTCAATGTATTCATGTGAATTCACCTTGGAAATGACACGGCTCAA	4438
OY	5008	CTTGTGTACATGGGATTAATGAAGAATTTTATAGCTCTCTAATATGGCGTACTGCAA	5067
Db	4439	CTTGTGTACATGGGATTAATGAAGAATTTTATAGCTCTCTAATATGGCGTACTGCAA	4498
OY	5068	GACCTCTTGAACACTTTCACAGAGATAGATATTTTAAGTCATGCGCTTGGCGTGTCTAT	5127
Db	4499	GACCTCTTGAACACTTTCACAGAGATAGATATTTTAAGTCATGCGCTTGGCGTGTCTAT	4558
OY	5128	GGCACCTTTCCTTCTGAAAGTCTGTGTTCTGTCCCAATGACCTTGGCCTTGTGAGCCGA	5187
Db	4559	GGCACCTTTCCTTCTGAAAGTCTGTGTTCTGTCCCAATGACCTTGGCCTTGTGAGCCGA	4618
OY	5188	GATGTGACCTGTGATTAAGGCGCAAAAGAGAGGCTGGGGCTTCTTCCCTCATCTGAAGAG	5247
Db	4619	GATGTGACCTGTGATTAAGGCGCAAAAGAGAGGCTGGGGCTTCTTCCCTCATCTGAAGAG	4678
OY	5248	CCCTTATTTGAATTCACGTGTGTGAGGCTTAGCCCTTCATTTCTGCACATTTCCCAACCTC	5307
Db	4679	CCCTTATTTGAATTCACGTGTGTGAGGCTTAGCCCTTCATTTCTGCACATTTCCCAACCTC	4738
OY	5308	CCAGCCCCCTTCAAGCAGACTAGGTGCTGTCAATTCACCCAAAGTGGGATTTGGCTTC	5367
Db	4739	CCAGCCCCCTTCAAGCAGACTAGGTGCTGTCAATTCACCCAAAGTGGGATTTGGCTTC	4797
OY	5368	CTTAAAGCTGGCTATTGTGCACATCCAGCAATACAGTTGGCTGCAAGAGCACACAGTG	5427
Db	4798	CTTAAAGCTGGCTATTGTGCACATCCAGCAATACAGTTGGCTGCAAGAGCACACAGTG	4857

```

QY 5428 GCCATTTTCCTTCACTGAGGGCTCAAACTCTGGACAAGTTGCTGCTCTGAGACCA 5487
DB 4858 GCCATTTTCCTTCACTGAGGGCTCAAACTCTGGACAAGTTGCTGCTCTGAGACCA 4917
QY 5488 GATATTTCCGAGAGCTGCTCAGTAAGGGGCCCGCTGAGGAACCTGCTCTTTTC 5547
DB 4918 GATATTTCCGAGAGCTGCTCAGTAAGGGGCCCGCTGAGGAACCTGCTCTTTTC 4977
QY 5548 TTTAAAGCCAGGCCCTCACTTACATTAACATTTTCAGGGTCACTGGAAAACATGAAGTGC 5607
DB 4978 TTTAAAGCCAGGCCCTCACTTACATTAACATTTTCAGGGTCACTGGAAAACATGAAGTGC 5037
QY 5608 CATTTTGAAGCTTACTGATGCGCCGCTCTCATTCACGTGTGCTGCTCATGCTCA 5667
DB 5038 CATTTTGAAGCTTACTGATGCGCCGCTCTCATTCACGTGTGCTGCTCATGCTCA 5097
QY 5668 CGAGAGAGCCAGCGCATGCGAGTGTCTCAATAGCTGTGCTCATTTGACAGAGAGGA 5727
DB 5098 CGAGAGAGCCAGCGCATGCGAGTGTCTCAATAGCTGTGCTCATTTGACAGAGAGGA 5157
QY 5728 AAGGTCACAAGAAAGTCAACTGGACAAGCAAGCCACCGGACATGCTGCTGTA 5787
DB 5158 AAGGTCACAAGAAAGTCAACTGGACAAGCAAGCCACCGGACATGCTGCTGTA 5217
QY 5788 AGGTTCAGACTGTGTGTGATCTGCAAGTCTTCACTGGAATTAATTATTCATTG 5847
DB 5218 AGGTTCAGACTGTGTGTGATCTGCAAGTCTTCACTGGAATTAATTATTCATTG 5277
QY 5848 CAGATCTTTTAGTGGGCAATTTATTCATTTCTGCTTTTAATTAACAAATGACCA 5907
DB 5278 CAGATCTTTTAGTGGGCAATTTATTCATTTCTGCTTTTAATTAACAAATGACCA 5337
QY 5908 AAAAACAAGTATCAAGCTGTTTAAGTCTTGGCTACTTGTCCCTGCTTCAAGTAGGC 5967
DB 5338 AAAAACAAGTATCAAGCTGTTTAAGTCTTGGCTACTTGTCCCTGCTTCAAGTAGGC 5397
QY 5968 CCGGTTTCCGAGTGTGATCTGTAAGGCTCAGCATGGGCTCAGCATGCTGCTTA 6027
DB 5398 CCGGTTTCCGAGTGTGATCTGTAAGGCTCAGCATGGGCTCAGCATGCTGCTTA 5457
QY 6028 ATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6087
DB 5458 ATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5517
QY 6088 GCGGTCACACTGTGTGTAAGCAGATGTTTGTCCGGAATTAATTAATGCTTGG 6143
DB 5518 GCGGTCACACTGTGTGTAAGCAGATGTTTGTCCGGAATTAATTAATGCTTGG 5573

```

RESULT 7
AC105902/c 16315 bp DNA linear PRI 27-JUN-2002
LOCUS Homo sapiens chromosome 3 clone RP11-578F5, complete sequence.
DEFINITION AC105902
ACCESSION AC105902.2 GI:21617714
VERSION HTG.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 16315)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.
and Haugen, E.D.
Direct Submission
Unpublished
2 (bases 1 to 16315)
Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.
Direct Submission
Submitted (10-JUN-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
3 (bases 1 to 16315)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,

COMMENT

Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C. and Haugen, E.D.
Direct Submission
Submitted (27-JUN-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
On Jun 27, 2002 this sequence version replaced gi:18104827.

Center: University of Washington Genome Center
Center Code: UWGC
Web site: <http://www.genome.washington.edu>
Contact: uwgc@u.washington.edu
Project Information

Center project name: chr-3
Center clone name: RP11-578F5 (bc0477)

Summary Statistics
Sequencing vector: plasmid, 100% of reads
Chemistry: Dye-terminator ET, 29% of reads
Chemistry: Dye-terminator Big Dye, 71% of reads
Assembly program: Phrap, version 0.990319
Consensus quality: 163063 bases at least Q40
Consensus quality: 163110 bases at least Q30
Consensus quality: 163115 bases at least Q20
Insert size: 163115; sum-of-contigs
Quality coverage: 9.1x in Q20 bases; sum-of-contigs

Overlapping Sequences:
5': RP11-68104 (UWGC:bc0528) AC104165
3': RP11-487J21 AC010170, 50065-bp overlap

Sequence Quality Assessment:
This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the Genbank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:
all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

ECORI	HindIII	BglII
SeqDerMap Fingerprint	SeqDerMap Fingerprint	SeqDerMap Fingerprint
-----	-----	-----
8696	8823	3091
-----	-----	-----
6	<800	6382
-----	-----	-----
740	789	512
-----	-----	-----
10200	10022	449
-----	-----	-----
1644	1614	5235
-----	-----	-----
		5065
-----	-----	-----
		1052
-----	-----	-----
		1030

	9393	9315	2845	2873	85529
	871	897	7321	7256	1207
	797	789	5620	5480	18057
	15170	15087	1720	1661	20661
	919	897	3901	4210	3557
	2182	2206	9970	9927	7796
	214	<800	3602	3542	1844
	6473	6579	522	<800	1354
	1456	1431	4155	4210	356
	682	<800	3891	3866	297
	3675	3703	3267	3284	9866
	3439	3449	50	<800	2935
	10009	10022	4248	4210	3034
	6068	6125	1035	1039	265
	11400	11202	3002	2986	6084
	96	<800	13316	13421	1766
	4100	4108	2978	2986	5507
	779	789	8146	8119	144
	1221	1181	4465	4507	1455
	278	<800	2239	2257	76
	8458	8511	45	<800	3696
	12014	11848	4368	4507	1021
	645	<800	286	<800	511
	1177	1181	4633	4819	1907
	153	<800	1462	1437	512
	1993	1940	2058	2083	367
	3153	3167	745	765	7100
	132	<800	1321	1310	5896
	4939	4944	3528	3542	1814
	466	<800	1640	1661	122
	1311	1295	8779	8805	43
	3183	3167	5089	5065	111
	16952	17153	2571	2605	1955
	1629	1614	2440	2393	3477
	7261	7351	1682	1661	4895
	7840	7927	1660	1661	774

Query Match	Best Local Match	Similarity	99.8%	Pred. No. 0	Mismatches	0	Indels	1	Gaps	1
Db	2409	CAACATCAATACAGATGAGATGCGGAGGAGCCCAAAAGTTTCAGAAAGGCGAAAGGACAA	74158	CAACATCAATACAGATGAGATGCGGAGGAGCCCAAAAGTTTCAGAAAGGCGAAAGGACAA	74099					
Qy	2469	TGACTCCCAATGTGTATTCAGATCATCCAGACACCAATGTATATGGCATCTGTACAGAA	2468							
Db	74098	TGACTCCCAATGTGTATTCAGATCATCCAGACACCAATGTATATGGCATCTGTACAGAA	74039							
Qy	2529	TTCCAGGCGCTCCTCTCCAGAGCGAGGATGAGACACCTACCCGCGTTCCAGAGGACACAT	2588							
Db	74038	TTCCAGGCGCTCCTCTCCAGAGCGAGGATGAGACACCTACCCGCGTTCCAGAGGACACAT	73979							
Qy	2589	GGGGGTCTGTCTCTCCCTCCACCCACCAACATATGCTCCAGGGCCCAACTGCAAGTTGAC	2648							
Db	73978	GGGGGTCTGTCTCTCCCTCCACCCACCAACATATGCTCCAGGGCCCAACTGCAAGTTGAC	73919							
Qy	2649	CACGTGAAGACCACTCTCTGCTCCCTCCTGAGTCTGAGAGTGAACCGTACACCTTCTC	2708							
Db	73918	CACGTGAAGACCACTCTCTGCTCCCTCCTGAGTCTGAGAGTGAACCGTACACCTTCTC	73855							
Qy	2709	CCATCCCAAAATGGGATGTAAAGCAGCAAGACACAGACATTCCTTACTGAACACTCA	2768							
Db	73858	CCATCCCAAAATGGGATGTAAAGCAGCAAGACACAGACATTCCTTACTGAACACTCA	73799							
Qy	2769	GGAGCCCATGAGAGCGACAGATTAACCTTGCATTCCTCAGACGCTTGTGAGTTCAAT	2828							
Db	73798	GGAGCCCATGAGAGCGACAGATTAACCTTGCATTCCTCAGACGCTTGTGAGTTCAAT	73739							
Qy	2829	AAGCAGGCGACTGAGACACCGGTCCGGTTCCTTAACAGAAATCCCTAAAGAAAGGAAT	2888							
Db	73738	AAGCAGGCGACTGAGACACCGGTCCGGTTCCTTAACAGAAATCCCTAAAGAAAGGAAT	73679							
Qy	2889	ATAAGAAAGAAACGACGAGAGTTTCTGTGACACCGCCCACTTGCATTTGCTCACTGA	2948							
Db	73678	ATAAGAAAGAAACGACGAGAGTTTCTGTGACACCGCCCACTTGCATTTGCTCACTGA	73619							

QY 2949 CTCATTCTAAGGCAAGACATTGAAATGATGATTCGATCTGATACAGTCATGACG 3008
| | | | |
Db 73618 CTCATTCTAAGGCAAGACATTGAAATGATGATTCGATCTGATACAGTCATGACG 73559
| | | | |
QY 3009 CTCATGTCCTCCAACTTAAAGGCTGTCGAGTTAGCCAGCCTGTAATGAGAGGAGAGGAC 3068
| | | | |
Db 73558 CTCATGTCCTCCAACTTAAAGGCTGTCGAGTTAGCCAGCCTGTAATGAGAGGAGAGGAC 73499
| | | | |
QY 3069 CTGAGTCACCTAGACATAGGGTTTGACAGCAAGCCCTGATTCAGAGTGTGTAACAGAGGCTT 3128
| | | | |
Db 73498 CTGAGTCACCTAGACATAGGGTTTGACAGCAAGCCCTGATTCAGAGTGTGTAACAGAGGCTT 73439
| | | | |
QY 3129 GCCCTTTTGAAGACAAAGTTCCAAAGAGGCTTACCTGAGGTCTCTACTCTCACT 3188
| | | | |
Db 73438 GCCCTTTTGAAGACAAAGTTCCAAAGAGGCTTACCTGAGGTCTCTACTCTCACT 73379
| | | | |
QY 3189 GGGGGTCCCAAGAGTGAAGAAAGCAATGTCCTTTTATTATTATTATTATTGGTGGTCCG 3248
| | | | |
Db 73378 GGGGGTCCCAAGAGTGAAGAAAGCAATGTCCTTTTATTATTATTATTATTGGTGGTCCG 73319
| | | | |
QY 3249 TGTATTATTAAGAGATCAAAATGATTAACAACCTAGCTCTTTTCACTGACTAGTATTAAC 3308
| | | | |
Db 73318 TGTATTATTAAGAGATCAAAATGATTAACAACCTAGCTCTTTTCACTGACTAGTATTAAC 73259
| | | | |
QY 3309 TCATACTAACTGCTTGGATGCTGGGTTGTGACTTCTACTGACCCGCTAGATTAAGCTGT 3368
| | | | |
Db 73258 TCATACTAACTGCTTGGATGCTGGGTTGTGACTTCTACTGACCCGCTAGATTAAGCTGT 73199
| | | | |
QY 3369 GCGTGTCCCAAGGTCGAGGTAATTTAACAATCTGTCCAACCAAGAAAGATGTCGT 3428
| | | | |
Db 73198 GCGTGTCCCAAGGTCGAGGTAATTTAACAATCTGTCCAACCAAGAAAGATGTCGT 73139
| | | | |
QY 3429 GTTTGAGCAGCATTTGACATATCTGCTTTGATTAAGAGACTTCTGATTTCTAGTCCG 3488
| | | | |
Db 73138 GTTTGAGCAGCATTTGACATATCTGCTTTGATTAAGAGACTTCTGATTTCTAGTCCG 73079
| | | | |
QY 3489 TTGCTGTGTTATCCCATTTGAGAAATTCATCTTGAATCCCATTTGCTATAGTCTTAGCAA 3548
| | | | |
Db 73078 TTGCTGTGTTATCCCATTTGAGAAATTCATCTTGAATCCCATTTGCTATAGTCTTAGCAA 73019
| | | | |
QY 3549 TAGAGAAATTTCTCAAGTTTCCATGTGCGGTTCTCTAGCGCAGCAATCTTTGACA 3608
| | | | |
Db 73018 TAGAGAAATTTCTCAAGTTTCCATGTGCGGTTCTCTAGCGCAGCAATCTTTGACA 72959
| | | | |
QY 3609 TTTTAAAGAGAAATTTAGAGAAATTTCTCATCTCTTAAATGTTTAAATATATACCAAC 3668
| | | | |
Db 72958 TTTTAAAGAGAAATTTAGAGAAATTTCTCATCTCTTAAATGTTTAAATATATACCAAC 72899
| | | | |
QY 3669 AGTGGCCCCCTGATTAAGTTTCTGTTGCACTGCAACCCATTACTTGGTAGCTTAAGAA 3728
| | | | |
Db 72898 AGTGGCCCCCTGATTAAGTTTCTGTTGCACTGCAACCCATTACTTGGTAGCTTAAGAA 72839
| | | | |
QY 3729 CAACCATTAAGCTTAAGTCTGCGGGAATGAGAAATTCAGAAATGAGATGCCCTGAATGAAA 3788
| | | | |
Db 72838 CAACCATTAAGCTTAAGTCTGCGGGAATGAGAAATTCAGAAATGAGATGCCCTGAATGAAA 72779
| | | | |
QY 3789 ATCAAAGTGTCAAGAGCTGTGCTCTCTCTGAAGGCTCTAGGAGAGAACCGGTTCTTG 3848
| | | | |
Db 72778 ATCAAAGTGTCAAGAGCTGTGCTCTCTCTGAAGGCTCTAGGAGAGAACCGGTTCTTG 72719
| | | | |
QY 3849 CCAATTTCAAGCTTCTAGAGGCTGTGCTCAATCCAGGCTTCCAGTGTGTCAGGCTTTT 3908
| | | | |
Db 72718 CCAATTTCAAGCTTCTAGAGGCTGTGCTCAATCCAGGCTTCCAGTGTGTCAGGCTTTT 72659
| | | | |
QY 3909 CTCACATGSCATACCTGTGACCTGCGCCCTCCACATTCCTCTTTGAATTAACAAACCCA 3968
| | | | |
Db 72658 CTCACATGSCATACCTGTGACCTGCGCCCTCCACATTCCTCTTTGAATTAACAAACCCA 72599
| | | | |
QY 3969 CCAGGAAGATCCAGAGTAATCTCTCCATCTAAAGATCTTCAATCATCTGAGGAAGGCTT 4028
| | | | |
Db 72598 CCAGGAAGATCCAGAGTAATCTCTCCATCTAAAGATCTTCAATCATCTGAGGAAGGCTT 72539
| | | | |
QY 4029 TTGCAATGCAAGCAACATAGCCAGGTGGGATTAGAACCAAGACATCTTTGGGGTGC 4088
| | | | |

Db 72538 TTGCAATGCAAGCAACATAGCCAGGTGGGATTAGAACCAAGACATCTTTGGGGTGC 72479
| | | | |
QY 4089 TGTATTCTGCTTACCAACCTTCTGCACTGACTCCACAGAGAGGCTCAAAATGA 4148
| | | | |
Db 72478 TGTATTCTGCTTACCAACCTTCTGCACTGACTCCACAGAGAGGCTCAAAATGA 72419
| | | | |
QY 4149 TCTGCGCAACAGGATGTTTGTGTTAGCTTGGGACTCTACACTT-AAAAAACCCAG 4207
| | | | |
Db 72418 TCTGCGCAACAGGATGTTTGTGTTAGCTTGGGACTCTACACTT-AAAAAACCCAG 72359
| | | | |
QY 4208 ATCAGAAATCTGSCATGCTGGGCTCAATTCACCTAGCAACAATGCTGGAGCT 4267
| | | | |
Db 72358 ATCAGAAATCTGSCATGCTGGGCTCAATTCACCTAGCAACAATGCTGGAGCT 72299
| | | | |
QY 4268 GGGCACACAGCTCTGCTTTTGAAGGGGTCCTCACTTCAACAGTCAACCAACCAACT 4327
| | | | |
Db 72298 GGGCACACAGCTCTGCTTTTGAAGGGGTCCTCACTTCAACAGTCAACCAACCAACT 72239
| | | | |
QY 4328 AGCCCTATCACTTCCCAATGAGGCTAAGTGTGTTTCTACTGATCAATGCCCTGC 4387
| | | | |
Db 72238 AGCCCTATCACTTCCCAATGAGGCTAAGTGTGTTTCTACTGATCAATGCCCTGC 72179
| | | | |
QY 4388 AGTTGCAATTAATGTAATGAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 4447
| | | | |
Db 72178 AGTTGCAATTAATGTAATGAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 72119
| | | | |
QY 4448 ACCATGAGCAACAATGTGCTCACATTAACCTTTCTTTTCTTTTCTTTTCTTTCTT 4507
| | | | |
Db 72118 ACCATGAGCAACAATGTGCTCACATTAACCTTTCTTTTCTTTTCTTTTCTTTCTT 72059
| | | | |
QY 4508 TTTTATTTTATGAGAGACAGATCTATCTGTTGCTTGGCTAGGCTGAGTGAAGGCGCA 4567
| | | | |
Db 72058 TTTTATTTTATGAGAGACAGATCTATCTGTTGCTTGGCTAGGCTGAGTGAAGGCGCA 71999
| | | | |
QY 4568 ATCTGGCTCACTGCAACTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4627
| | | | |
Db 71998 ATCTGGCTCACTGCAACTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 71939
| | | | |
QY 4628 CAAATAGCTGGGATCACTGCGCAAAACACATGCGCAGCTAATTTTGTGTTGTTGTTG 4687
| | | | |
Db 71938 CAAATAGCTGGGATCACTGCGCAAAACACATGCGCAGCTAATTTTGTGTTGTTGTTGTTG 71879
| | | | |
QY 4688 AGACAGGTTTCAACCATGTTGCGCAGGCTGTCTCAACCTCTGCGGCTCAAGCAATCTCTC 4747
| | | | |
Db 71878 AGACAGGTTTCAACCATGTTGCGCAGGCTGTCTCAACCTCTGCGGCTCAAGCAATCTCTC 71819
| | | | |
QY 4748 CTGCTCGGCTCCCAAGTGTGAGATTACAGATGTGAGCCACCGCATCCAGCCCCACA 4807
| | | | |
Db 71818 CTGCTCGGCTCCCAAGTGTGAGATTACAGATGTGAGCCACCGCATCCAGCCCCACA 71759
| | | | |
QY 4808 CCTCATTTATCACAATTAACCTGCGCAGCTGTAACCTGTGGAATTTTGTCTCTCAACCCCTGCT 4867
| | | | |
Db 71758 CCTCATTTATCACAATTAACCTGCGCAGCTGTAACCTGTGGAATTTTGTCTCTCTCAACCCCTGCT 71699
| | | | |
QY 4868 CTGATCTGGAAGAGAGGATTAATGTTATGCTTGTGACAGACAGTCCCAAGTTCAATATT 4927
| | | | |
Db 71698 CTGATCTGGAAGAGAGGATTAATGTTATGCTTGTGACAGACAGTCCCAAGTTCAATATT 71639
| | | | |
QY 4928 TCTGCGGCAAAACCTTCTTCAAAATTAATGTAATCTTCAATGTAATCAATGTAATCAAC 4987
| | | | |
Db 71638 TCTGCGGCAAAACCTTCTTCAAAATTAATGTAATCTTCAATGTAATCAATGTAATCAAC 71579
| | | | |
QY 4988 TTGGAATGACCGGCTCAACTGTTTCAATGAGCAATGAGCAATGAGAAAGAAATTTATAGTCTC 5047
| | | | |
Db 71578 TTGGAATGACCGGCTCAACTGTTTCAATGAGCAATGAGCAATGAGAAAGAAATTTATAGTCTC 71519
| | | | |
QY 5048 CTAAATGAGGCTGTACTGCAAGACTCTTGAACAATTCAGAGATAGAGATATTAAAGTC 5107
| | | | |
Db 71518 CTAAATGAGGCTGTACTGCAAGACTCTTGAACAATTCAGAGATAGAGATATTAAAGTC 71459
| | | | |
QY 5108 ATGCCCTTGGGCTTGCCTATGAGCACTTTCCCTTCTGAAAGTCTGCTTCCCAAGTGA 5167
| | | | |

Db	71458	ATGCGCTTGGCGTTGCTGACTGACACCTTTCCCTTCGAAAGTGTGGTTCCTGCCAGTGA	71339
Qy	5168	CCCTTGGCCCTTGTGAGCCGAGATGCTGACCCCTGCTATTAAGGGCCAAAGAGGGCTCGGCC	5227
Db	71398	CCCTTGGCCCTTGTGAGCCGAGATGCTGACCCCTGCTATTAAGGGCCAAAGAGGGCTCGGCC	71339
Qy	5228	TTCCCTTCCCTCACTGAAGAGCCCTTATTTGAATTCACTGTTGTGAGACCCCTAGCCCTCAT	5287
Db	71338	TTCCCTTCCCTCACTGAAGAGCCCTTATTTGAATTCACTGTTGTGAGACCCCTAGCCCTCAT	71279
Qy	5288	TCTCGACATTTCCCAACCTCCAGCCCTTCCAGACAGACTAGTGGCCCTGATTCAC	5347
Db	71278	TCTCGACATTTCCCAACCTCCAGCCCTTCCAGACAGACTAGTGGCCCTGATTCAC	71219
Qy	5348	CCAAAGTGGGATTTGGCCCTTCTTAAGGTGGCTACTTGACACTACACGACATCACTGTT	5407
Db	71218	CCAAAGTGGGATTTGGCCCTTCTTAAGGTGGCTACTTGACACTACACGACATCACTGTT	71159
Qy	5408	GCCCTGCAAGGACACACAGTGGCCATTTTCCCTCAACTGAAGGGCTCAAACTCCTGGACAA	5467
Db	71158	GCCCTGCAAGGACACACAGTGGCCATTTTCCCTCAACTGAAGGGCTCAAACTCCTGGACAA	71099
Qy	5468	GTTTGCCTGGCTTCCTGAGACCACTATTCTCTGGAGCTGCTCAGTGAAGGGCCACGCT	5527
Db	71098	GTTTGCCTGGCTTCCTGAGACCACTATTCTCTGGAGCTGCTCAGTGAAGGGCCACGCT	71039
Qy	5528	GAGGAACCCCTGGGCTTTTCTTTAAAGCCGAGGCCCACTTAATPAAACATTTTCAGGGT	5587
Db	71038	GAGGAACCCCTGGGCTTTTCTTTAAAGCCGAGGCCCACTTAATPAAACATTTTCAGGGT	70979
Qy	5588	CAC TGGAACAGTGAAGTGCATTTGTGTAAGCTACTGACATGCCACGCCACTGCTCATC	5647
Db	70978	CAC TGGAACAGTGAAGTGCATTTGTGTAAGCTACTGACATGCCACGCCACTGCTCATC	70919
Qy	5648	CACGTGCTTCGCCATGCTCTACGAGAAGGCGCACGSCATGCAAGACTGTGTCTTAATGCTG	5707
Db	70918	CACGTGCTTCGCCATGCTCTACGAGAAGGCGCACGSCATGCAAGACTGTGTCTTAATGCTG	70859
Qy	5708	TGGCTATTGGACAGAAAGGAAAGGTCTCAAGGAAAGATCAACTGGGACAAGACAAGAGCC	5767
Db	70858	TGGCTATTGGACAGAAAGGAAAGGTCTCAAGGAAAGATCAACTGGGACAAGACAAGAGCC	70739
Qy	5768	ACCGGACATGAGCCCTTGGTAAAGGTTAGACAGACTGTGTGTGTGGATCTGCACTGCTCAC	5827
Db	70798	ACCGGACATGAGCCCTTGGTAAAGGTTAGACAGACTGTGTGTGTGGATCTGCACTGCTCAC	70739
Qy	5828	TGGAAATAATTATTCATTTGCATATCTTTTAGTGGCATTTTATTCATTTTCTGTGCT	5887
Db	70738	TGGAAATAATTATTCATTTGCATATCTTTTAGTGGCATTTTATTCATTTTCTGTGCT	70679
Qy	5888	TTAAATTAACAAATGTATCCAAAAAAACAAGATACACTGTGTTAAAGCTTCGGCTACTTG	5947
Db	70678	TTAAATTAACAAATGTATCCAAAAAAACAAGATACACTGTGTTAAAGCTTCGGCTACTTG	70619
Qy	5948	TCCCCTGTGTTCAGTAGAGGCCCGGCTTCCCACTGTGTACTGTGACAGGCTTCAGCATGG	6007
Db	70618	TCCCCTGTGTTCAGTAGAGGCCCGGCTTCCCACTGTGTACTGTGACAGGCTTCAGCATGG	70559
Qy	6008	GCTCAGACAGATGCGTCTTAATTTTGGAGATGATACAGAAAGCCAGGCTTGGGATPACAG	6067
Db	70558	GCTCAGACAGATGCGTCTTAATTTTGGAGATGATACAGAAAGCCAGGCTTGGGATPACAG	70499
Qy	6068	TTCTTTTCTCTTCATTTGATGCGGTGACACTGTGTGAAGCAGATGTTTTGTCCGAAATA	6127
Db	70498	TTCTTTTCTCTTCATTTGATGCGGTGACACTGTGTGAAGCAGATGTTTTGTCCGAAATA	70439
Qy	6128	AAAAATAATAGCTTTGAGATCTGGCCA 6153	
Db	70438	AAAAATAATAGCTTTGAGATCTGGCCA 70413	

LOCUS AC104165 200956 bp DNA linear PRI 27-JUN-2002
DEFINITION Homo sapiens chromosome 3 clone RPL1-68104, complete sequence.
ACCESSION AC104165 AC068625
VERSION AC104165.2 GI:21617715
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 200956)
Kaul,R.K., Olson,M.V., Zhou Y., James,R.A., Rouse,G., Wu,Z.,
Saenpimmachak,C., Phelps,K.A., Buckley,D., Kitukawa,M., Raymond,C.
and Haugen,E.D.
Direct Submission
Unpublished
TITLE Direct Submission
JOURNAL
AUTHORS 2 (bases 1 to 200956)
Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.
Direct Submission
Submitted (05-DEC-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
JOURNAL
AUTHORS 3 (bases 1 to 200956)
Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenpimmachak,C., Phelps,K.A., Buckley,D., Kitukawa,M., Raymond,C.
and Haugen,E.D.
Direct Submission
Submitted (27-JUN-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
COMMENT On Jun 27, 2002 this sequence version replaced gi:17352429.

Sequencing vector: unknown; 42% of reads
 Sequencing vector: plasmid; 58% of reads
 Chemistry: Dye-terminator Ed; 80% of reads
 Chemistry: Dye-terminator Big Dye; 20% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 200732 bases at least Q40
 Consensus quality: 200909 bases at least Q30
 Consensus quality: 200947 bases at least Q20
 Insert size: 200556; sum-of-contents
 Quality coverage: 10.4x in Q20 bases; sum-of-contents

Overlapping Sequences:

5': RP11-348P10 (UMGC:bc0377) AC124045
3': RP11-578F5 (UMGC:bc0477) AC105902, 95352-bp overlap

Sequence Quality Assessment: This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the Genbank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:
This sequence has been validated by Multiple Complete Digest

fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

BgIII		EcoRI		HindIII	
SeqDerMap	FgPrnt	SeqDerMap	FgPrnt	SeqDerMap	FgPrnt
5328	5288	8696	8571	3534	3560
2067	2057	6	<800	6382	6318
6846	6934	6068	6154	512	<800
144	<800	11400	11207	449	<800
1459	1439	96	<800	5736	5646
70	<800	4100	4106	2978	3031
3690	3725	779	760	8146	8133
102	<800	1221	1170	4465	4477
511	<800	278	<800	2239	2257
1903	1919	8458	8571	45	<800
5124	5028	12014	11879	4368	4477
367	<800	645	<800	286	<800
7100	6934	1177	1170	4633	4784
5890	5889	153	<800	1462	1405
1814	1816	1993	1932	2058	2110
125	<800	3153	3186	745	<800
43	<800	132	<800	1321	1243
117	<800	4939	4965	3528	3560
1955	1919	466	<800	1640	1662
3477	3529	1311	1279	8779	8765
4899	5028	3183	3186	5089	4991
774	799	16952	16678	2571	2631
10749	10616	1629	1629	2440	2446
1620	1602	7261	7427	1682	1662
8868	8810	7837	7939	1660	1662
5884	5889	3940	4106	4912	4991
1521	1529	756	760	2470	2446
771	799	2395	2496	8892	8765
11706	11675	5310	5399	1257	1243

369	<800	21139	21498	10114	9869
5177	5288	4082	4106	14019	14339
2916	2950	7866	7939	3580	3560
137	<800	7411	7427	3022	3031
2406	2347	3378	3366	8296	8133
3857	3725	7272	7427	3319	3311
5249	5288	3059	3186	1694	1662
5343	5288	8214	8571	1747	1662
6819	6934	5836	5849	2370	2446
2335	2347	5394	5399	4817	4784
569	<800	419	<800	1139	1073
135	<800	3260	3186	3946	3909
3825	3725	2082	2097	12931	12902
6741	6934	8568	8571	2109	2110
1472	1439	1681	1629	364	<800
3707	3725	741	760	78	<800
703	<800	2905	3005	254	<800
207	<800	6129	6318	1674	1662
5047	5028	1134	1133	1250	1243
1070	1069	4675	4675	1083	1073
4665	4675	1121	1133	1023	1073
6920	6934	2602	2631	2449	2446
3317	3264	1683	1697	1249	1243
3525	3529	7228	7182		

Query Match 53.0%; Score 3268; DB 9; Length 200956;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 3728; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY	2409	CAACATCAATCTAGATGATGCGGACGACCAAAAGTTTCAGAAAGGCGAAAGACAA	2468
DB	179863	CAACATCAATCTAGATGATGCGGACGACCAAAAGTTTCAGAAAGGCGAAAGACAA	179804
QY	2469	TGACTCCCATGTGTATGACGATCATCGAGACACATGTATATGGGCATCTGCTACAGA	2528
DB	179803	TGACTCCCATGTGTATGACGATCATCGAGACACATGTATATGGGCATCTGCTACAGA	179744
QY	2529	TTCCAGCGGCTCTCTCTGACAGCAGAGGTGACACTACCGGCGCTTCAGGACCACT	2588
DB	179743	TTCCAGCGGCTCTCTCTGACAGCAGAGGTGACACTACCGGCGCTTCAGGACCACT	179684
QY	2589	GGGGTCTGTCTCTCCCTCCCAACCACTATGCTTCAGGGCCCAACTGCAAAGTTGGC	2648

Db 179683 GGGGGCTGTCTCTCCCTCCCAACCAATATGCTCCAGAGGCCCACTGCAAGTTGCG 179624
Qy 2649 CACTAGAGAGCCACCTCTGCTCCCTCTGAGTCTGAGATGAGACCGTACACCTTCTC 2708
Db 179683 CACTAGAGAGCCACCTCTGCTCCCTCTGAGTCTGAGATGAGACCGTACACCTTCTC 179664
Qy 2709 CCATCCCAACAAATGGGGATGTAAAGACAGAGACAGACATTTCCCTTACTGAAACCTCA 2768
Db 179563 CCATCCCAACAAATGGGGATGTAAAGACAGAGACAGACATTTCCCTTACTGAAACCTCA 179504
Qy 2769 GAGAGCCATGAGAGCCAGCAAGATTAATCTGATCTCATTTCCAGAGCTTTGCTGATTTATA 2828
Db 179503 GAGAGCCATGAGAGCCAGCAAGATTAATCTGATCTCATTTCCAGAGCTTTGCTGATTTATA 179444
Qy 2829 AAGAGGGGCACTGAGACACCGCTGCTTCTTAACCAAGATTCCTAAAGAGAGGAATT 2888
Db 179443 AAGAGGGGCACTGAGACACCGCTGCTTCTTAACCAAGATTCCTAAAGAGAGGAATT 179384
Qy 2889 ATACAGAGAGAACAGAGAGAGTTTCTGAGACACCGCCAACTTGACATTTGCTCAGTGA 2948
Db 179383 ATACAGAGAGAACAGAGAGAGTTTCTGAGACACCGCCAACTTGACATTTGCTCAGTGA 179324
Qy 2949 CTCAATCTTAAGGGCAAGACATTTGAAATGATGAAATTTCCATCTGATACATGATGACAG 3008
Db 179323 CTCAATCTTAAGGGCAAGACATTTGAAATGATGAAATTTCCATCTGATACATGATGACAG 179264
Qy 3009 CTCAATGCTCTCTCAACTTAAAGGCTGAGGCTGAGCAGCTGATTAAGAGAGAGAGGC 3068
Db 179263 CTCAATGCTCTCTCAACTTAAAGGCTGAGGCTGAGCAGCTGATTAAGAGAGAGAGGC 179204
Qy 3069 CTGAGTCACTTACATTAAGGGTTGACAGACAGCCCTGATTCAGAGTGTAAACAGAGGCTT 3128
Db 179203 CTGAGTCACTTACATTAAGGGTTGACAGACAGCCCTGATTCAGAGTGTAAACAGAGGCTT 179144
Qy 3129 GCGCTCTTACAGACAAAGTCCCAATTTCCAGAGGCTTACCTGAGTCTTACTCTACT 3188
Db 179143 GCGCTCTTACAGACAAAGTCCCAATTTCCAGAGGCTTACCTGAGTCTTACTCTACT 179084
Qy 3189 GGGGCTCCAGGATGAAAAAGCAATGTCCTTTTATTAATTAATTTGGTGGTCCG 3248
Db 179083 GGGGCTCCAGGATGAAAAAGCAATGTCCTTTTATTAATTAATTTGGTGGTCCG 179024
Qy 3249 TGTATTTTAAAGATCAATATGATAACCACTAGCTCTTTTCACTGACTTAATAAAC 3308
Db 179023 TGTATTTTAAAGATCAATATGATAACCACTAGCTCTTTTCACTGACTTAATAAAC 178964
Qy 3309 TCATACTAACTGTTTGAATGCTGAGTTGTGATCTTCTAATGACCGCTAGATTAACGTGT 3368
Db 178963 TCATACTAACTGTTTGAATGCTGAGTTGTGATCTTCTAATGACCGCTAGATTAACGTGT 178904
Qy 3369 GCGTGTCCCGAGGTGGGGAATTTTAAATCTGTTCACACGAAAGAAATGTGT 3428
Db 178903 GCGTGTCCCGAGGTGGGGAATTTTAAATCTGTTCACACGAAAGAAATGTGT 178844
Qy 3429 GTTTGAGCAGATTAACATATGCTTGTGATAAGAGACTTCTGATTTCTAGTGGG 3488
Db 178843 GTTTGAGCAGATTAACATATGCTTGTGATAAGAGACTTCTGATTTCTAGTGGG 178784
Qy 3489 TTTGTTGTTATCCATGTGAAATTCATCTTGAATCCCATTTGCTATATGCTTGAACA 3548
Db 178783 TTTGTTGTTATCCATGTGAAATTCATCTTGAATCCCATTTGCTATATGCTTGAACA 178724
Qy 3549 TTAAGGAAATTTCCCAAGTTTCCATGTCGGTTCTCTTACGTCGAGCAATCTTTGACA 3608
Db 178723 TTAAGGAAATTTCCCAAGTTTCCATGTCGGTTCTCTTACGTCGAGCAATCTTTGACA 178664
Qy 3609 TTTAAGAGAAATTTAGAGAAATTTCTCATCTCTTAAATGTTTAAATATATACCAAC 3668
Db 178663 TTTAAGAGAAATTTAGAGAAATTTCTCATCTCTTAAATGTTTAAATATATACCAAC 178604
Qy 3669 AGTGGCCCCCTGCATTAATGTTTCTGTGCACTGCAACCCATTAATCTGTGATCTTAATAA 3728
Db 178603 AGTGGCCCCCTGCATTAATGTTTCTGTGCACTGCAACCCATTAATCTGTGATCTTAATAA 178544

Qy 3729 CAACATTTAGCTTATAGTCTGAGGATCAGAAATTCAAAATGATGTCCCTGAATGAA 3788
Db 178543 CAACATTTAGCTTATAGTCTGAGGATCAGAAATTCAAAATGATGTCCCTGAATGAA 178484
Qy 3789 ATCAAGGTGTGAGCAGAGCTGTGCTCTTCTGAAAGCTTAAAGAGAACCGGTTCTTG 3848
Db 178483 ATCAAGGTGTGAGCAGAGCTGTGCTCTTCTGAAAGCTTAAAGAGAACCGGTTCTTG 178424
Qy 3849 CCATTTCAAGCTTCTAAGAGCTGTGATTTCCAGAGCTCCAGTGTCTGTCAACTTTT 3908
Db 178423 CCATTTCAAGCTTCTAAGAGCTGTGATTTCCAGAGCTCCAGTGTCTGTCAACTTTT 178364
Qy 3909 CTCAATGAGCATCACTGTGACACTGAGCCCTCCCACTTCCCTTTGACTTAAGAGCCA 3968
Db 178363 CTCAATGAGCATCACTGTGACACTGAGCCCTCCCACTTCCCTTTGACTTAAGAGCCA 178304
Qy 3969 CCAGAGAGATCCAGATTAATCTCTCATCTAAAGATCTTCAATCTCTGAAAGCCTT 4028
Db 178303 CCAGAGAGATCCAGATTAATCTCTCATCTAAAGATCTTCAATCTCTGAAAGCCTT 178244
Qy 4029 TTGCAATGCAAGACAAATAGCCACAGGTGGGATTTAGACAGACATCTTTGGGGTGC 4088
Db 178243 TTGCAATGCAAGACAAATAGCCACAGGTGGGATTTAGACAGACATCTTTGGGGTGC 178184
Qy 4089 TGTATTTGCTCTACACACCTTCTGCACTGACCTCCACAGAGAGGCTCAAAATGA 4148
Db 178183 TGTATTTGCTCTACACACCTTCTGCACTGACCTCCACAGAGAGGCTCAAAATGA 178124
Qy 4149 TCTGCGCAGACAGAGATGTTTGTGTTAGCTTGGGACCTTAACACTT-AAAAAACCCGAG 4207
Db 178123 TCTGCGCAGACAGAGATGTTTGTGTTAGCTTGGGACCTTAACACTT-AAAAAACCCGAG 178064
Qy 4208 ATCAGAAATCTGGCATGCTGTGGGCTCAATTTCTACCTAGACAACTGTGCTGAGGT 4267
Db 178063 ATCAGAAATCTGGCATGCTGTGGGCTCAATTTCTACCTAGACAACTGTGCTGAGGT 178004
Qy 4268 GGGCAGCAGCTCTGCTTTTAAAGGGGTGTCCACTTCAACAGGTACCAACAGCCCACT 4327
Db 178003 GGGCAGCAGCTCTGCTTTTAAAGGGGTGTCCACTTCAACAGGTACCAACAGCCCACT 177944
Qy 4328 AGCCCTTATCACTTCCCAATAGAGCTTAAGTGTGTTTCTTATCTGATCAATGCCCTGC 4387
Db 177943 AGCCCTTATCACTTCCCAATAGAGCTTAAGTGTGTTTCTTATCTGATCAATGCCCTGC 177884
Qy 4388 AGTTGCATTTATTTGTAATGAAAAAGAAAGACTGGGATTAATCTTAATCAGGTAGTAC 4447
Db 177883 AGTTGCATTTATTTGTAATGAAAAAGAAAGACTGGGATTAATCTTAATCAGGTAGTAC 177824
Qy 4448 ACCATGAGACCAATGAGTGTCAATTAACCTTTTCTTTTCTTTTCTTTTCTTTTCTT 4507
Db 177823 ACCATGAGACCAATGAGTGTCAATTAACCTTTTCTTTTCTTTTCTTTTCTTTTCTT 177764
Qy 4508 TTTTCTTTTAAATGTAAGACAGGATCTAATCTGTTGCTAGGCTGAGTGGGCGCA 4567
Db 177763 TTTTCTTTTAAATGTAAGACAGGATCTAATCTGTTGCTAGGCTGAGTGGGCGCA 177704
Qy 4568 ATCTGGCTCACTGCAACTTGTCTCTCTGAGCTCAAGCAATTTCTCCACTCAAGCTTCC 4627
Db 177703 ATCTGGCTCACTGCAACTTGTCTCTCTGAGCTCAAGCAATTTCTCCACTCAAGCTTCC 177644
Qy 4628 CAAATAGCTGGGATACAGTGGCAAAACAGCAACATGCGAGCTTAATTTGATTTTGTAG 4687
Db 177643 CAAATAGCTGGGATACAGTGGCAAAACAGCAACATGCGAGCTTAATTTGATTTTGTAG 177584
Qy 4688 AGACAGGGTTTCAACATGTTGCCAGGCTGTCTCAACTCTCTGGGCTCAAGCAATCTCTC 4747
Db 177583 AGACAGGGTTTCAACATGTTGCCAGGCTGTCTCAACTCTCTGGGCTCAAGCAATCTCTC 177524
Qy 4748 CTGCTCGGCTCCCAAGTGTGGGATTTAAGATGTGAGCCACCGCATCCAGCCCCACA 4807
Db 177523 CTGCTCGGCTCCCAAGTGTGGGATTTAAGATGTGAGCCACCGCATCCAGCCCCACA 177464

QY 4808 CCTCATTTATACCAATTACCTGCCCCAGTACCTGAGCTTTTGGTCTTCTCACCCCTGCT 4867
 DB 177463 CCTCATTTATACCAATTACCTGCCCCAGTACCTGAGCTTTTGGTCTTCTCACCCCTGCT 177404
 QY 4868 CTGATGTGGAAGAGAGAGGATTAATGTTATAGCTTGCAGACAGCTGCCAAGTTTCAATAT 4927
 DB 177403 CTGATGTGGAAGAGAGAGGATTAATGTTATAGCTTGCAGACAGCTGCCAAGTTTCAATAT 177344
 QY 4928 TCTGCGGCAAAAATTCTTCAAAAATAATGTAATTTGATTTGATTTCAATGATTTCACT 4987
 DB 177343 TCTGCGGCAAAAATTCTTCAAAAATAATGTAATTTGATTTGATTTCAATGATTTCACT 177284
 QY 4988 TTGGAATGACACCGCTCTCACTTTGTCATGAGCATTAATGAAGAATTTTATAGCTTC 5047
 DB 177283 TTGGAATGACACCGCTCTCACTTTGTCATGAGCATTAATGAAGAATTTTATAGCTTC 177224
 QY 5048 CTAAATGGGCTGACTGCAAGACCTCTGAAACCTTCCAGAGGATAGGATTTTAAATGTC 5107
 DB 177223 CTAAATGGGCTGACTGCAAGACCTCTGAAACCTTCCAGAGGATAGGATTTTAAATGTC 177164
 QY 5108 ATGCCCCCTGGCGTTGCTATGACACCTTTCCCTTCTGAAAGTCTGCTGCCCCAGTGA 5167
 DB 177163 ATGCCCCCTGGCGTTGCTATGACACCTTTCCCTTCTGAAAGTCTGCTGCCCCAGTGA 177104
 QY 5168 CCTTGGCTTGTGAGCCGAGATGCTGACCTGTCATTAAGGCGCAAGAGGCGTGGCG 5227
 DB 177103 CCTTGGCTTGTGAGCCGAGATGCTGACCTGTCATTAAGGCGCAAGAGGCGTGGCG 177044
 QY 5228 TTCCCTTCCCTCACTGAAGACCCCTTATTTGAATTCATCTGAGGAGCCCTGAGCCCTCAT 5287
 DB 177043 TTCCCTTCCCTCACTGAAGACCCCTTATTTGAATTCATCTGAGGAGCCCTGAGCCCTCAT 176984
 QY 5288 TCTCGACATTTCCCACTCTCCAGCCCTTCCCAAGAGAGACTAGTACCCTGATTTCCAG 5347
 DB 176983 TCTCGACATTTCCCACTCTCCAGCCCTTCCCAAGAGAGACTAGTACCCTGATTTCCAG 176924
 QY 5348 CCAAGGTGGATTTGGCCTTCTTAAAGCTGAGCTACTTGTCACTACCAATCACTGCTT 5407
 DB 176923 CCAAGGTGGATTTGGCCTTCTTAAAGCTGAGCTACTTGTCACTACCAATCACTGCTT 176864
 QY 5408 GCCTGCAAGGACACACGCTGGCCATTTCTTCACTGAGGAGCTCAAACTCTCGAGCA 5467
 DB 176863 GCCTGCAAGGACACACGCTGGCCATTTCTTCACTGAGGAGCTCAAACTCTCGAGCA 176804
 QY 5468 GTTGTGCTGCTCTGAGACCAAGTATTTCTGAGAGCTGCTCAGTGAAGGCGCCAGCT 5527
 DB 176803 GTTGTGCTGCTCTGAGACCAAGTATTTCTGAGAGCTGCTCAGTGAAGGCGCCAGCT 176744
 QY 5528 GAGGAACCTGCTTTTCTTTTAAAGCCAGGCCCCCACTTACATTAACATTTCAAGGT 5587
 DB 176743 GAGGAACCTGCTTTTCTTTTAAAGCCAGGCCCCCACTTACATTAACATTTCAAGGT 176684
 QY 5588 CACTGGAACAGTGAAGTGCATTTGTAAGCTGCTGAGTGCAGGCCCACTGCTCATC 5647
 DB 176683 CACTGGAACAGTGAAGTGCATTTGTAAGCTGCTGAGTGCAGGCCCACTGCTCATC 176624
 QY 5648 CACTGCTGCTGCAATGCTTACGAGAGAGCCAGCGATGAGGATGCTCTAATGCTG 5707
 DB 176623 CACTGCTGCTGCAATGCTTACGAGAGAGCCAGCGATGAGGATGCTCTAATGCTG 176564
 QY 5708 TGTCTATTGACAGAGAGAGAGTCTCAAGAGAGTCACTGGGACAGACAGCAAGCCC 5767
 DB 176563 TGTCTATTGACAGAGAGAGAGTCTCAAGAGAGTCACTGGGACAGACAGCAAGCCC 176504
 QY 5768 ACCGAGCATGGCTTGTAAAGGTGTAGACAGCTGCTGAGTGTGAGATCTGAGTGTGCTAC 5827
 DB 176503 ACCGAGCATGGCTTGTAAAGGTGTAGACAGCTGCTGAGTGTGAGATCTGAGTGTGCTAC 176444
 QY 5828 TGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5887
 DB 176443 TGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 176384
 QY 5888 TTAATTAACAAATTAATTAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5947

DB 176383 TTAATTAACAAATTAATTAACAAATTAATTAATTAATTAATTAATTAATTAATTAAT 176324
 QY 5948 TCCCTGCTGCTAGTAGAGCCCGGCTTCCAGTTTGTGACTGTGACAGGCTCAGCATG 6007
 DB 176323 TCCCTGCTGCTAGTAGAGCCCGGCTTCCAGTTTGTGACTGTGACAGGCTCAGCATG 176264
 QY 6008 GCTCAGCAGATGCTGCTTAAATTTGTGATGATGATGATGATGATGATGATGATGATGATG 6067
 DB 176263 GCTCAGCAGATGCTGCTTAAATTTGTGATGATGATGATGATGATGATGATGATGATG 176204
 QY 6068 TTCTTCCCTTCAATTTGATGCGCTGCACTGTGTGAAGAGATGTTTTGTCCGGAATA 6127
 DB 176203 TTCTTCCCTTCAATTTGATGCGCTGCACTGTGTGAAGAGATGTTTTGTCCGGAATA 176144
 QY 6128 AAAATTAATGCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 6153
 DB 176143 AAAATTAATGCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 176118

RESULT 9
 AK026622
 LOCUS 3224 bp mRNA linear PRI 29-SBP-2000
 DEFINITION Homo sapiens CDNA: FLJ22969 fls, clone KAT10759.
 ACCESSION AK026622
 VERSION AK026622.1 GI:10439514
 KEYWORDS oligo cloning; fls (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
 NEDO human cDNA sequencing project
 TITLE Unpublished
 JOURNAL 2 (bases 1 to 3224)
 REFERENCE Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
 NEDO human cDNA sequencing project
 AUTHORS Shibahara,T., Tanaka,T. and Nakamura,Y.
 TITLE Direct Submission

COMMENT
 JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)
 NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES
 source
 1..3224
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="KAT10759"
 /cell_line="KATO III"
 /cell_type="signet-ring cell carcinoma"
 /clone_1fb="KAT"
 /note="Cloning vector pMB18SFL3"
 111..2621
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAB1511.1"
 /db_xref="GI:10439515"
 /translation="MAGLNCVSIALLGLGAARLPGEAEFIALPRESNITVLI KLGPETLAKCYIVISKRHTLMSIKSGERIVFPSCSPENHVEIQRKNDGSG PCPRGEVLOQSTSLPLTNLFTIMDVAAHSIGLEQPSIPRLQIDGSGSCPGSVT HSIQRIDATVVRIGTFCSSNGTVSRIRKQEGVRAHLHPFRPRVSGSILNRSIK RLCTIISVFEDEGSATLMSANYPEGFPEDELTWQFVVAHLRASVPLNLSNCR

KEERVEYIIPGSTINPEVFKLEDPGPMAGNFILSGCDODASPEILRLQROVLV
QHPOINENKIIYVVDLSNERAMSLTIEPPVQSRKFRVGCCELESRTCSNLTLSG
SKHISFLCDDLTRLMMANVERKTSIEPTDYQKRSYSLQVSDILHLEHLEPSMKL
LVPRDLSTLVPAOKLOQTHREKSCPSYLVASALPSODLYGSGCPGSIKOIO
VKONISVTLRTPAPSPROASROGLTVEFIEYKEBGEVPLTPDKRSVYLRTPMDR
GLPSTISVSNMISVPRDOVACTIFPKESGVVCTGRAPMIIQERTABEETPSLDE
VLPRSPFHHSFWNISCSPTSGKQDLRPSVLTPTPTVDLTVLAAVGSVLLS
ALGLICCVKKKKKKTKNGPAPVGINGINTEMPQKPFQKGRKDNDSHYAVIEDT
MYGHLLODSSGSLQPEVDYTRPFGTGMGCPSPPTICRPAKPLATEBPPRSP
PESSEPYTFHPNNGDVSKDITDILSTQEPMEPAP"

BASE COUNT 789 a 922 c 810 g 703 t
ORIGIN

48.7%; Score 3004; DB 9; Length 3224;

Query Match 99.9%; Pred. No. 0;
Best Local Similarity 0; Mismatches 4; Indels 0; Gaps 0;
Matches 3204; Conservative

173 GCGCGCAGTGTAGTACGCGAGGCGGAGCGAGCTGCCCGGGCTTGGGCGCTGGGGCC 232
1 GCGCGCAGTGTAGTACGCGAGGCGGAGCGAGCTGCCCGGGCTTGGGCGCTGGGGCC 60
233 GCGCGTCCCAACCGTGTTCCTCCACGAGGCGGAGGCGTCCGGAGTCATGCGCGGCC 292
61 GCGCGTCCCAACCGTGTTCCTCCACGAGGCGGAGGCGTCCGGAGTCATGCGCGGCC 120
293 TGAACCTGGGGGTCTCTATCGCACTGTAGGGGTTCGTGCTGGGGTGGCGCGCGCTGC 352
121 TGAACCTGGGGGTCTCTATCGCACTGTAGGGGTTCGTGCTGGGGTGGCGCGCGCTGC 180
353 GCGCGGCGGCGAGAGCTTTGAGATTGCTCTGCCAGAGAAAGCAACATTACAGTTCTCA 412
181 GCGCGGCGGCGAGAGCTTTGAGATTGCTCTGCCAGAGAAAGCAACATTACAGTTCTCA 240
413 TAAAGCTGGGAGCCCGGACTGTGCTGGGAAACCGTGTACATCTCTATTCTAAAGAC 472
241 TAAAGCTGGGAGCCCGGACTGTGCTGGGAAACCGTGTACATCTCTATTCTAAAGAC 300
473 ATATACCATGTGTCTCATCAAGTCTGGAGAAAGATAGTCTTTACCTTTAGCTGCCAGA 532
301 ATATACCATGTGTCTCATCAAGTCTGGAGAAAGATAGTCTTTACCTTTAGCTGCCAGA 360
533 GTCTGAGAAATCATCTTTGTATAGAGATCCAGAAAAATATGACTGTATGTCAGGCCAT 592
361 GTCTGAGAAATCATCTTTGTATAGAGATCCAGAAAAATATGACTGTATGTCAGGCCAT 420
593 GTCTTTTGGGGAGGTTCAAGCTTCAGCCCTCGACATCGTGTGGCTACCTTCACAGAA 652
421 GTCTTTTGGGGAGGTTCAAGCTTCAGCCCTCGACATCGTGTGGCTACCTTCACAGAA 480
653 CTTTCATCTGGGATGTCAAAAGCTCATTAAGAGCATCGTTTAGAGCTGAGTTTTCATCC 712
481 CTTTCATCTGGGATGTCAAAAGCTCATTAAGAGCATCGTTTAGAGCTGAGTTTTCATCC 540
713 CTGCGCTGAGGAGATCGGTCCGGGTGAGAGCTGCCAGACGAGATCATCTCCATCA 772
541 CTGCGCTGAGGAGATCGGTCCGGGTGAGAGCTGCCAGACGAGATCATCTCCATCA 600
773 GCGGCGGAATGATCTCCACCGTGTGAGATCGGAACTTTGCGAGCAATGGCACTGTGT 832
601 GCGGCGGAATGATCTCCACCGTGTGAGATCGGAACTTTGCGAGCAATGGCACTGTGT 660
833 CCGGATCAAGATGCAAGAGGAGTGAATAATGGCTTACCTCCCATGTTTCCACCCCA 892
661 CCGGATCAAGATGCAAGAGGAGTGAATAATGGCTTACCTCCCATGTTTCCACCCCA 720
893 GAAATGTCTCGGCTTTCAGCATTTGCAACCGCTCATCTATAAAGCTGTGTGCATCATCG 952
721 GAAATGTCTCGGCTTTCAGCATTTGCAACCGCTCATCTATAAAGCTGTGTGCATCATCG 780
953 AGTCTGTGTGAGGAGTGAAGGCTCAGCAACCTGTATGTCTGCCAATACCAAGAGGCT 1012
781 AGTCTGTGTGAGGAGTGAAGGCTCAGCAACCTGTATGTCTGCCAATACCAAGAGGCT 840

1013 TCCCTGAGGATGAGCTCATGACGTGCGAGTTTGTCTTCTGCGACACTCGCGGCGCACCG 1072
841 TCCCTGAGGATGAGCTCATGACGTGCGAGTTTGTCTTCTGCGACACTCGCGGCGCACCG 900
1073 TCTCCTTCTCAACTTCAACCTTCCACTGTGAGAGGAGAGGAGGAGGAGGAGTGAATCT 1132
901 TCTCCTTCTCAACTTCAACCTTCCACTGTGAGAGGAGAGGAGGAGGAGGAGTGAATCT 960
1133 ACATCCCGGGCTCCACACCAACCCCGAGTGTTCACCTGAGAGGAGAGGAGGAGGAGGAG 1192
961 ACATCCCGGGCTCCACACCAACCCCGAGTGTTCACCTGAGAGGAGAGGAGGAGGAGGAG 1020
1193 ACATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1252
1021 ACATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
1253 GATCTCTCGGCTGCGAGTTTCCAGTTTGTGTTCACATCCACAAATGAAAGCAATAAA 1312
1081 GATCTCTCGGCTGCGAGTTTCCAGTTTGTGTTCACATCCACAAATGAAAGCAATAAA 1140
1313 TCTACGTGTGATCTTGAATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1372
1141 TCTACGTGTGATCTTGAATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
1373 TCAACAGAGCGGCAAGTTTGTCCCTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1432
1201 TCAACAGAGCGGCAAGTTTGTCCCTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
1433 GTAGCAACCTCACTGTGATCTGTGCTCCAAACACAAATCTCTTCTTGTGTGTGTGT 1492
1261 GTAGCAACCTCACTGTGATCTGTGCTCCAAACACAAATCTCTTCTTGTGTGTGTGT 1320
1493 TGACACGTGTGTGATGTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1552
1321 TGACACGTGTGTGATGTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
1553 AAAGGAAATCTTCACTCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1612
1381 AAAGGAAATCTTCACTCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
1613 ATGACTTCTCTGGAAGCTGT 1672
1441 ATGACTTCTCTGGAAGCTGT 1500
1673 CCGAGAGCTGAGGAGATCAACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1732
1501 CCGAGAGCTGAGGAGATCAACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
1733 TGCCAGTGCATACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1792
1561 TGCCAGTGCATACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
1793 TCAAGCATCAAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1852
1621 TCAAGCATCAAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680
1853 TCCAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1912
1681 TCCAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
1913 AAGGCGTTTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1972
1741 AAGGCGTTTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800
1973 GAGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2032
1801 GAGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1860
2033 AGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2092
1861 AGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1920
2093 CATTCATGATCATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2152

Db	1921	CATTCAATGATCATCCAGGACGAGCGGACCCCGGGCTGAGGAATCTTCAGCTCGACGAGG	1980
Qy	2153	ATGAGCTCCCCAAGCAAGCTTCCACCATCAAGATTCTGGGTCAACATCTCTAACTGCA	2212
Db	1981	ATGAGCTCCCCAAGCAAGCTTCCACCATCAAGCTTCTGGGTCAACATCTCTAACTGCA	2040
Qy	2213	GCCCCAGAGCGGCAAGCAGCTAGACCTGCTCTTCTCGGTGACACTTAACCCAGAGACTG	2272
Db	2041	GCCCCAGAGCGGCAAGCAGCTAGACCTGCTCTTCTCGGTGACACTTAACCCAGAGACTG	2100
Qy	2273	TGACATTGACTGTCATCTCTCATTCGACGGGTGGAGGTGGAGTCTTACTGCTGTCTGCC	2332
Db	2101	TGACATTGACTGTCATCTCTCATTCGACGGGTGGAGGTGGAGTCTTACTGCTGTCTGCC	2160
Qy	2333	TCGGGCTCATATTGCTGTGTGAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG	2392
Db	2161	TCGGGCTCATATTGCTGTGTGAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG	2220
Qy	2393	TGGGTATCTCAATGCGAACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	2452
Db	2221	TGGGTATCTCAATGCGAACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	2280
Qy	2453	AAGGGCGAAAGAGCAATGCACTCCCATGCTGTATGAGTCATGTAGAGACACCAATGATATAG	2512
Db	2281	AAGGGCGAAAGAGCAATGCACTCCCATGCTGTATGAGTCATGTAGAGACACCAATGATATAG	2340
Qy	2513	GGCATCTGCTCAACAGGATTCGACGGGCTCTTCTCTGACGCGAGAGTGGACACTACCGGC	2572
Db	2341	GGCATCTGCTCAACAGGATTCGACGGGCTCTTCTCTGACGCGAGAGTGGACACTACCGGC	2400
Qy	2573	CGTTCGAGGCAACATGGGGGTCTGTCTCTCCCTCCCAACCAATATGCTTCAGGGCCC	2632
Db	2401	CGTTCGAGGCAACATGGGGGTCTGTCTCTCCCTCCCAACCAATATGCTTCAGGGCCC	2460
Qy	2633	CAACTGCAAAATTGAGCCACTAGAGAGCCACTCGTCTCCCTCCCTCTGAGCTGAGAGTG	2692
Db	2461	CAACTGCAAAATTGAGCCACTAGAGAGCCACTCGTCTCCCTCTGAGCTGAGAGTG	2520
Qy	2693	AACGGTACACCTTCTCCCATCCCAACAAATGGGAGTGTAAAGCAGACAGACACATTC	2752
Db	2521	AACGGTACACCTTCTCCCATCCCAACAAATGGGAGTGTAAAGCAGACAGACACATTC	2580
Qy	2753	CCTTACTGAACATCTCAGAGGCCATGTAGCCAGACAGATTACTTATCATTTCCAGACGC	2812
Db	2581	CCTTACTGAACATCTCAGAGGCCATGTAGCCAGACAGATTACTTATCATTTCCAGACGC	2640
Qy	2813	TTTGTCTGAGTTCTTAAAGCAGGGCACTGAGACACCCGTCCGTGTCTCTTAACCAAGATC	2872
Db	2641	TTTGTCTGAGTTCTTAAAGCAGGGCACTGAGACACCCGTCCGTGTCTCTTAACCAAGATC	2700
Qy	2873	CTTAAAGAAAGGAATTATACAGAAAGAAACAGCAGAGGTTTTTCTGACACCGCAACTT	2932
Db	2701	CTTAAAGAAAGGAATTATACAGAAAGAAACAGCAGAGGTTTTTCTGACACCGCAACTT	2760
Qy	2933	CACATTGTCTAGTGAACCTCATTTCTTAAGGGCAAGACATTGAAATGATGAATTTCCAACTG	2992
Db	2761	CACATTGTCTAGTGAACCTCATTTCTTAAGGGCAAGACATTGAAATGATGAATTTCCAACTG	2820
Qy	2993	GATACAGTCATGACAGCTCATGTGCTCTCAACTTAAGGCTGTGCGGTTAGCCAGCCTGTA	3052
Db	2821	GATACAGTCATGACAGCTCATGTGCTCTCAACTTAAGGCTGTGCGGTTAGCCAGCCTGTA	2880
Qy	3053	ATGAGAGAGAGAGAGCCGTGATCTACACTAGCATTAAGGTTTGGACGAAAGCCCTGGAATTCAG	3112
Db	2881	ATGAGAGAGAGAGAGCCGTGATCTACACTAGCATTAAGGTTTGGACGAAAGCCCTGGAATTCAG	2940
Qy	3113	TGTTAAACAGAGGCTTGCCCTCTTACAGGACAAAGTTTCAATTCCAAAGACCTTACCTGA	3172
Db	2941	TGTTAAACAGAGGCTTGCCCTCTTACAGGACAAAGTTTCAATTCCAAAGACCTTACCTGA	3000
Qy	3173	GGTCCCTTACTCTCACTGGGGTCCCGAGATGAAAACGAATGTGCTTTTATTTATAT	3232

Db	3001	GGGTCCTTCACTCACTGGGGGCTCCCGAGGAAGAAACGACAAATGCGCTTTTATATATAT	3060
Qy	3233	TTATTGGTGGTCTGCTGTTATTTAAGAGATCAATATATTAACCACTAGCTCTTTTCAC	3292
Db	3061	TTATTGGTGGTCTGCTGTTATTTAAGAGATCAATATATTAACCACTAGCTCTTTTCAC	3120
Qy	3293	CTGACTTAGTAATTAACATCACTACTGCTGTTGGATGCGGGTGTGACTTCTACTGAC	3352
Db	3121	CTGACTTAGTAATTAACATCACTACTGCTGTTGGATGCGGGTGTGACTTCTACTGAC	3180
Qy	3353	CGCTAGATAACGTCGTGCTGCTCCCA	3380
Db	3181	CGCTAGATAACGTCGTGCTGCTCCCA	3208
RESULT 10			
LOCUS	AK026028	2113 bp	mRNA linear PRI 29-SEP-2000
DEFINITION	Homo sapiens cDNA: FLJ22375 fis, clone HRC06785.		
ACCESSION	AK026028		
VERSION	AK026028.1	GI:10438737	
KEYWORDS	oligo capping; fis (full insert sequence).		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (sites)		
	Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.		
TITLE	NEDO human cDNA sequencing project		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2113)		
AUTHORS	Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-VUG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-Ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)		
COMMENT	NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry for Biotechnology; cDNA library construction; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing; Departent of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).		
FEATURES	Location/Qualifiers		
SOURCE	1..2113		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="HRC06785"		
	/cell_type="Primary human renal epithelial cells"		
	/clone_lib="HRC"		
	/note="Cloning vector pME18SFL3"		
BASE COUNT	585 a 538 c 470 g 520 t		
ORIGIN			
	Query Match 27.6%; Score 1703; DB 9; Length 2113;		
	Best Local Similarity 99.9%; Pred. No. 0;		
	Matches 1753; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
Qy	2448	TCGAGAAAGGGCGAAGACATATGATCTCCCATGTGTATATGATGATCGACGACACCATGCT	2507
Db	1	TCGAGAAAGGGCGAAGACATATGATCTCCCATGTGTATATGATGATCGACGACACCATGCT	60
Qy	2508	ATATGGGCACTGCTGACAGATTCCAGCGGCTCCTTCTGTCAGCGAGGTGACACCTTA	2567
Db	61	ATATGGGCACTGCTGACAGATTCCAGCGGCTCCTTCTGTCAGCGAGGTGACACCTTA	120
Qy	2568	CGGCGCTTCCAGGGACCAATGGGGGCTGTGCTCCTCCCTCCCAACCAACATATATGCTTCAG	2627

Db 121 CCGGCGGCTCCAGGGCCACCATGGGGGTCTGTCCTCCCTCCACCCACCATATGCTCCAG 180
Qy 2628 GGGCCCACTGCAAAATGGGSCACATGAGAGGACCTCTGCTCCCTCCAGATCTGA 2687
Db 181 GGGCCCACTGCAAAATGGGSCACATGAGAGGACCTCTGCTCCCTCCAGATCTGA 240
Qy 2688 GAGTGAACCGTACACCTTCTCCCATCCCAACAATGGGATGTAGCAGCAGAGACAGA 2747
Db 241 GAGTGAACCGTACACCTTCTCCCATCCCAACAATGGGATGTAGCAGCAGAGACAGA 300
Qy 2748 CATTCCCTTACTGAAACATCTCAGAGGCCATGAGCCAGCAAAATTAATTGATCCATTCCA 2807
Db 301 CATTCCCTTACTGAAACATCTCAGAGGCCATGAGCCAGCAAAATTAATTGATCCATTCCA 360
Qy 2808 GACGCTTGTGAGTGTCTTAAGAAGAGGCACTGAGACACCCGTCGGTCTTAACCAAG 2867
Db 361 GACGCTTGTGAGTGTCTTAAGAAGAGGCACTGAGACACCCGTCGGTCTTAACCAAG 420
Qy 2868 AAATCTTAAAGAGGAATTTATAGAAAGAAACAGCAGAGAGTTCCTTGACACCGCC 2927
Db 421 AAATCTTAAAGAGGAATTTATAGAAAGAAACAGCAGAGAGTTCCTTGACACCGCC 480
Qy 2928 AACCTTCACTTGTCTCAGTGAATCTTAAGGGCAAGCATTGAAATGATGAATTCGA 2987
Db 481 AACCTTCACTTGTCTCAGTGAATCTTAAGGGCAAGCATTGAAATGATGAATTCGA 540
Qy 2988 ATCTGGAATACAGTCAATGACATGATGCTCTCACTTAGGCTGTGGGTTACCCAGC 3047
Db 541 ATCTGGAATACAGTCAATGACATGATGCTCTCACTTAGGCTGTGGGTTACCCAGC 600
Qy 3048 CTGTAAATGAGAGAGAGAGGCTGAGTCACTAGCATAGGTTGAGCAAGGCTTGATTT 3107
Db 601 CTGTAAATGAGAGAGAGAGGCTGAGTCACTAGCATAGGTTGAGCAAGGCTTGATTT 660
Qy 3108 CAGAGTGTAAACAGAGGCTTCCCTTTCAGAGCAACAGTTCAATTCAGAGAGCCTTA 3167
Db 661 CAGAGTGTAAACAGAGGCTTCCCTTTCAGAGCAACAGTTCAATTCAGAGAGCCTTA 720
Qy 3168 CCTGAGGTCCTCACTCTCACTGAGGTCCTCCAGAGATGAAACAGCAATGTGCTTTTATTT 3227
Db 721 CCTGAGGTCCTCACTCTCACTGAGGTCCTCCAGAGATGAAACAGCAATGTGCTTTTATTT 780
Qy 3228 ATTATTTATTTGATGATGCTGCTGATTTATTAAGAGATCAAAATGTATTAACCACTAGCTCT 3287
Db 781 ATTATTTATTTGATGATGCTGCTGATTTATTAAGAGATCAAAATGTATTAACCACTAGCTCT 840
Qy 3288 TTCACTGACTTAATTAATTAATCACTAATGATGATGATGATGATGATGATGATGATGAT 3347
Db 841 TTCACTGACTTAATTAATTAATCACTAATGATGATGATGATGATGATGATGATGATGAT 900
Qy 3348 CTGACCGCTGAGTAAACGTCGTGCTGCTCCCGAGGTGAGGAAATTAATTAATTAATTAAT 3407
Db 901 CTGACCGCTGAGTAAACGTCGTGCTGCTCCCGAGGTGAGGAAATTAATTAATTAATTAAT 960
Qy 3408 CAACCAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3467
Db 961 CAACCAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Qy 3468 CTTCTGATTTCTCTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3527
Db 1021 CTTCTGATTTCTCTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Qy 3528 ATTGTCTTAATGATCTTGAAGAAATTAAGAAATTTCTCAAGTTTCAATGTGCGGTTCTCT 3587
Db 1081 ATTGTCTTAATGATCTTGAAGAAATTAAGAAATTTCTCAAGTTTCAATGTGCGGTTCTCT 1140
Qy 3588 AGCTGAGCAATATCTTGAATTAAGAAATTTAGAAATTTAGAAATTTAGAAATTTAGAAATTT 3647
Db 1141 AGCTGAGCAATATCTTGAATTAAGAAATTTAGAAATTTAGAAATTTAGAAATTTAGAAATTT 1200
Qy 3648 ATGTTTAAATATATACAAACAGTGGCCCTGATTAATTTTCTGTTGCACTGCAAC 3707

Db 1201 ATGTTTAAATATATACAAACAGTGGCCCTGATTAATTTTCTGTTGCCACTGCAAC 1260
Qy 3708 CATTACTTGTAGTCTTAATAAACAACATTAATAGTCTGAGGATACGAATTTCCAA 3767
Db 1261 CATTACTTGTAGTCTTAATAAACAACATTAATAGTCTGAGGATACGAATTTCCAA 1320
Qy 3768 AATGATGTCTCTGATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 3827
Db 1321 AATGATGTCTCTGATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 1380
Qy 3828 TAGGAGAAAGCCGCTTCTGTCATTTCAAGCTTTAGAGGCTGCTGATTTCCAGGCT 3887
Db 1381 TAGGAGAAAGCCGCTTCTGTCATTTCAAGCTTTAGAGGCTGCTGATTTCCAGGCT 1440
Qy 3888 CCAAGTGTCTGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3947
Db 1441 CCAAGTGTCTGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Qy 3948 CTCTTTGACTTAACAAGCCCAACAGAGATTCAGAGATTAATCTCAATTAAGATCT 4007
Db 1501 CTCTTTGACTTAACAAGCCCAACAGAGATTCAGAGATTAATCTCAATTAAGATCT 1560
Qy 4008 TCATCATCTCTGAGAGACCTTTTGGCATGCAAGACATAGCCACAGGTGGGATTAGG 4067
Db 1561 TCATCATCTCTGAGAGACCTTTTGGCATGCAAGACATAGCCACAGGTGGGATTAGG 1620
Qy 4068 ACCAGACATCTTTGGGAGGCTTATCTGCTCAACACCTTCTGCACTGATCTCC 4127
Db 1621 ACCAGACATCTTTGGGAGGCTTATCTGCTCAACACCTTCTGCACTGATCTCC 1680
Qy 4128 ACAGAGAGGCTACAAATATGATCTGCGCACAGGATGTTTGTAGCTTGCGGATCT 4187
Db 1681 ACAGAGAGGCTACAAATATGATCTGCGCACAGGATGTTTGTAGCTTGCGGATCT 1740
Qy 4188 AACACTTAAAAAA 4201
Db 1741 AACACTTAAAAAA 1754

RESULT 11
AR243778 2209 bp DNA linear PAT 20-DEC-2002
LOCUS AR243778
DEFINITION Sequence 16 from patent US 6476195.
ACCESSION AR243778
VERSION AR243778.1 GI:27291271
KEYWORDS
SOURCE
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 2209)
AUTHORS Komatsu, S., Rosen, C.A., Ruben, S.M., Duan, R.D., Moore, P.A., Shi, Y., LaFleur, D.W., Wei, J., Florence, K.A., Young, P., Brewer, L.A., Soppel, D.R., Endress, G.A., Ebner, R., Olsen, H. and Muenster, M.
TITLE Secreted protein HNF20
JOURNAL Patent: US 6476195-A 16-05-NOV-2002;
FEATURES
source location/Qualifiers
1..2209 /organism="unknown"
BASE COUNT 508 a 646 c 566 g 477 t 12 others
ORIGIN

Query Match 25.1%; Score 1545; DB 6; Length 2209;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 2145; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 204 AGCTGCGCGGCTTGGGCGCTGAGGCGCGCTGCCACCGTGTTCCTCCACCGAG 263
Db 2 AGCTGCGCGGCTTGGGCGCTGAGGCGCGCTGCCACCGTGTTCCTCCACCGAG 61
Qy 264 GCGGAGGCGTCCCGGAGTATGCGCGGCTGAACTGCGGCGTCTTATGCACTGTAAG 323
Db 62 GCGGAGGCGTCCCGGAGTATGCGCGGCTGAACTGCGGCGTCTTATGCACTGTAAG 121

QY 324 GGTTCGTGCTGGGTGCGCGCGCTGCCGCGCGGGGCAAGACTTTTGAAGTTGCTCT 383
DB 122 GGTTCGTGCTGGGTGCGCGCGCTGCCGCGCGGGGCAAGACTTTTGAAGTTGCTCT 181
QY 384 GCCACGAGAAAGCAATTAACAGTTCTCAATTAAGCTGGGGACCCGCACTCTGTGGCAAA 443
DB 182 GCCACGAGAAAGCAATTAACAGTTCTCAATTAAGCTGGGGACCCGCACTCTGTGGCAAA 241
QY 444 ACCCTGTTACATCGTCAATTTCTAAAGACATATATACATGTTGTCCATCAAGCTGTGAGA 503
DB 242 ACCCTGTTACATCGTCAATTTCTAAAGACATATATACATGTTGTCCATCAAGCTGTGAGA 301
QY 504 AAGATAGTCTTTACCTTTAGCTGCCAGAGTCTTGAAATCACTTTGTCAATAGATTCOA 563
DB 302 AAGATAGTCTTTACCTTTAGCTGCCAGAGTCTTGAAATCACTTTGTCAATAGATTCOA 361
QY 564 GAAATATTTGACTGATGTCAGAGCCGCAATGCTCTTTGGGGAGAGTTCAAGTTCAAGCCCTC 623
DB 362 GAAATATTTGACTGATGTCAGAGCCGCAATGCTCTTTGGGGAGAGTTCAAGTTCAAGCCCTC 421
QY 624 GACATGTTGTTGCTACCTCAACAGAACTTTCACTGGAGTGTCAAAAGCTCATTAAG 683
DB 422 GACATGTTGTTGCTACCTCAACAGAACTTTCACTGGAGTGTCAAAAGCTCATTAAG 481
QY 684 CATGGTTTGAAGTSCAGTTTTCATCCCTGCTGAGGAGATCGTTCGGGTGAG 743
DB 482 CATGGTTTGAAGTSCAGTTTTCATCCCTGCTGAGGAGATCGTTCGGGTGAG 541
QY 744 CTGCCAGAGCGAGTCACTCACTCACTCAAGCGGCGCAATCAAGCCAGCTGTCAAGAT 803
DB 542 CTGCCAGAGCGAGTCACTCACTCACTCAAGCGGCGCAATCAAGCCAGCTGTCAAGAT 601
QY 804 CGGAACCTTCTGAGCAATGSCACTGTGTCCTGGATCAAGATGCAAGAAAGAGTGAAT 863
DB 602 TGGAACTTCTGAGCAATGSCACTGTGTCCTGGATCAAGATGCAAGAAAGAGTGAAT 661
QY 864 GGCCTTACACCTTCCATGTTCCACCAGAAATGTCCTCGGCTTCAAGATTGCAAAACCG 923
DB 662 GGCCTTACACCTTCCATGTTCCACCAGAAATGTCCTCGGCTTCAAGATTGCAAAACCG 721
QY 924 CTCATATATTAACCGTCTGTCATCAATGATGTCGTTGAGGAGTGAAGGCTCACCAAC 983
DB 722 CTCATATATTAACCGTCTGTCATCAATGATGTCGTTGAGGAGTGAAGGCTCACCAAC 781
QY 984 CCTGATGTCTGCAATACCCAGAAAGGCTTCCCTGAGATGAGTCAATGACGTGAGATT 1043
DB 782 CCTGATGTCTGCAATACCCAGAAAGGCTTCCCTGAGATGAGTCAATGACGTGAGATT 841
QY 1044 TGTGTTCTGTGACACCTGCGGGCCAGCGTCTCTTCTCAACTTCAACCTTCCAACTG 1103
DB 842 TGTGTTCTGTGACACCTGCGGGCCAGCGTCTCTTCTCAACTTCAACCTTCCAACTG 901
QY 1104 TGAAGAGAGAGAGACCGGTTGAATATCAATCCCGGGTCCACCAACCAACCCCGAGGT 1163
DB 902 TGAAGAGAGAGAGACCGGTTGAATATCAATCCCGGGTCCACCAACCAACCCCGAGGT 961
QY 1164 GTTCAAGCTGAGAGACAGAGCTGGGAAATGAGCGGGAACTTCAACCTCTCTGTGA 1223
DB 962 GTTCAAGCTGAGAGACAGAGCTGGGAAATGAGCGGGAACTTCAACCTCTCTGTGA 1021
QY 1224 AGGCTGTGACCAAGATGCCCAAGATCCAGAGATCTCCGGCTGACAGTTCCAAATTTTGGT 1283
DB 1022 AGGCTGTGACCAAGATGCCCAAGATCCAGAGATCTCCGGCTGACAGTTCCAAATTTTGGT 1081
QY 1284 CCAACATCCACAAATAAGAAAGCAATTAATCTACGTGTGTAAGTTGATATAGCGAGC 1343
DB 1082 CCAACATCCACAAATAAGAAAGCAATTAATCTACGTGTGTAAGTTGATATAGCGAGC 1141
QY 1344 CATGTCACTACCATGAGGACGCGCCGTCAAACAGAGCGCAAGTTGTCCTGAGCTG 1403
DB 1142 CATGTCACTACCATGAGGACGCGCCGTCAAACAGAGCGCAAGTTGTCCTGAGCTG 1201

QY 1404 TTTCGTGTCTTGAATCTCGGACCTGCAATGCAACTCACCTTGACATCTTGCTCCAA 1463
DB 1202 TTTCGTGTCTTGAATCTCGGACCTGCAATGCAACTCACCTTGACATCTTGCTCCAA 1261
QY 1464 ACACAAATCTCTCTCTTTGATGATCTGACAGTCTGTGGATGGAATGGAAAAAAC 1523
DB 1262 ACACAAATCTCTCTCTTTGATGATCTGACAGTCTGTGGATGGAATGGAAAAAAC 1321
QY 1524 CATTAAGCTGACAGACCAACCGGTACTGCGAAAGAAATCTTACTCACTCAGGTGCCAG 1583
DB 1322 CATTAAGTGCACAGACCAACCGGTACTGCGAAAGAAATCTTACTCACTCAGGTGCCAG 1381
QY 1584 TGAATCTCTCCACTGCTGTGTGAGTGTGATGATCTTCTCTGAAAGTCTGTGCCAA 1643
DB 1382 TGAATCTCTCCACTGCTGTGTGAGTGTGATGATCTTCTCTGAAAGTCTGTGCCAA 1441
QY 1644 GGAAGAGCTCAGCTGTGTGCTGTGTGCTGAGCCAGAGAGCTGACAGATCAACAGAGAA 1703
DB 1442 GGAAGAGCTCAGCTGTGTGCTGTGTGCTGAGCCAGAGAGCTGACAGATCAACAGAGAA 1501
QY 1704 GCCCTGCAACACCAAGCTTCAAGTCACTGCTGAGCCAGTGCATACCAAGCAGGACCTGTA 1763
DB 1502 GCCCTGCAACACCAAGCTTCAAGTCACTGCTGAGCCAGTGCATACCAAGCAGGACCTGTA 1561
QY 1764 CTTGGCTCTCTTCTGCGCGGAGGCTCTATCAAGCAATCCAGGTGAACAGAACTCTC 1823
DB 1562 CTTGGCTCTCTTCTGCGCGGAGGCTCTATCAAGCAATCCAGGTGAACAGAACTCTC 1621
QY 1824 GGTGACCTTGTGACCTTGTGCGCCAGCTTCAACAAAGAGGCTCCAGGACGAGGTGTGAC 1883
DB 1622 GGTGACCTTGTGACCTTGTGCGCCAGCTTCCGACAAAGAGGCTTCCAGGACGAGGTGTGAC 1681
QY 1884 GGTGTCTTTATACCTTATTTTAAAGAGAAAGCGTTTTCACAGTGAACCTTGACACAAA 1943
DB 1682 GGTGTCTTTATACCTTATTTTAAAGAGAAAGCGTTTTCACAGTGAACCTTGACACAAA 1741
QY 1944 AAGCAAGTCTTACCTGAGAGACCCCACTGGGACCGGGGCTGCCATCCTTCACTCTGT 2003
DB 1742 AAGCAAGTCTTACCTGAGAGACCCCACTGGGACCGGGGCTGCCATCCTTCACTCTGT 1801
QY 2004 GTTCTGGAACATCAGGTGCTCCAGAGACAGGATGAGCTGCTGACTTCTTTAAGGAGCG 2063
DB 1802 GTTCTGGAACATCAGGTGCTCCAGAGACAGGATGAGCTGCTGACTTCTTTAAGGAGCG 1861
QY 2064 GAGCGGCTGTGTCTGCAAGACAGGCGGCAATTCATGATCATCCAGAGACGAGACCGG 2123
DB 1862 GAGCGGCTGTGTCTGCAAGACAGGCGGCAATTCATGATCATCCAGAGACGAGACCGG 1921
QY 2124 GGTGAGAGATCTTCAAGCTGCTGAGAGAGATGTCTCCCAAGCCAAAGCTTCCACATTA 2183
DB 1922 GGTGAGAGATCTTCAAGCTGCTGAGAGAGATGTCTCCCAAGCCAAAGCTTCCACATTA 1981
QY 2184 CAGCTTCTGGGTCAACATCTCACTGACAGGCCCAAGAGGAGGCAAGCAGTGAAGCTGCT 2243
DB 1982 CAGCTTCTGGGTCAACATCTTAAATTAAGAGCCCAAGAGGAGGCAAGCAGTGAAGCTGCT 2041
QY 2244 CTTCTCGGTGACATTAACCCCAAGAGCTGTGACTTGAATCTGATCATCTCATGACAGCGT 2303
DB 2042 CTTCTCGGTGACATTAACCCCAAGAGCTGTGACTTGAATCTGATCATCTCATGACAGCGT 2101
QY 2304 GGAAGGTGAGTCTTAATCTGCTGTGCTGCTGCGGCTCATCATTTGTGTGTGAAAAA 2360
DB 2102 GGAAGGTGAGTCTTAATCTGCTGTGCTGCTGCGGCTCATCATTTGTGTGTGAAAAA 2158

RESULT 12
AK026187 1241 bp mRNA linear PRI 29-SEP-2000
LOCUS AK026187
DEFINITION Homo sapiens CDNA: FLJ22534 fls, clone HRC13020.
ACCESSION AK026187
VERSION AK026187.1 GI:10438956
KEYWORDS oligo capping, fls (full insert sequence).
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases)
AUTHORS Kawabata,A., Hkaji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shidhara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1241)
AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shidhara,T., Tanaka,T. and Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction: 5'- & 3'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
source location/Qualifiers
1..1241
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HRC13020"
/cell_type="Primary human renal epithelial cells"
/clone_1fb="HRC"
/note="Cloning vector pME18SFL3"
65..>1241
/note="unnamed protein product"
/codon_start=1
/protein_id="BA15318.1"
/db_xref="GI:10438957"
/translation="MAGNENLSLOGDDOASPGILRLQFOVNLHPQNESKIYVD
LSNERAMSLTIEPRPVKSRKVPFGFVLESRTSSNLTLSGSHKISFLCDLUTR
LMNVEKITSCTDHRVCKRKSYSLOVPSDILHPVLDHPSKILVVKORLSLVVPA
QRLQOHTHEKPCNTSFSYLVASAIPOSDYFGSFCPGSISKQIYVKONISVLTFTFA
SFOQASRQGLTFSFYPKEGEVFTPTDTSKYVLRTPMDRGLSITSVSMISV
PRDOVACLTFPKERSGVGVCOTGRAPMIIOERTABEISFLEDELTPRSFPHHSFMV
NISNSGPTSGKQDLDFSVTLTPRYVDLTIVIIAIVGGSVLISLGLITICVKKKKK
"
BASE COUNT 296 a 380 c 313 g 252 t
ORIGIN
Query Match 19.1%; Score 1177; DB 9; Length 1241;
Beet Local Similarity 99.9%; Pred. No. 0;
Matches 1227; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1133 ACATCCGGGCTCCACCAACCCCGAGGTGTTCAAGCTGAGGACAGCAAGCCTGGGA 1192
Db 3 ACATCCGGGCTCCACCAACCCCGAGGTGTTCAAGCTGAGGACAGCAAGCCTGGGA 62
QY 1193 ACATGGCGGGAACTTCAACTCTCTCTGCAAGAGCTGTGACCAAGTCCCAAGTCCAG 1252
Db 63 ACATGGCGGGAACTTCAACTCTCTCTGCAAGAGCTGTGACCAAGTCCCAAGTCCAG 122
QY 1253 GGATCTCGGGGTGAGTTCGAAGTTTGGTCCACATCCAAATGAAATGAACATTAATA 1312
Db 123 GGATCTCGGGGTGAGTTCGAAGTTTGGTCCACATCCAAATGAAATGAACATTAATA 182
QY 1313 TCTAGTGGTTGACTTGAATTAAGCGAGCATGTCACTCAATCGAGCCAGCGCCG 1372
Db 183 TCTAGTGGTTGACTTGAATTAAGCGAGCATGTCACTCAATCGAGCCAGCGCCG 242
QY 1373 TCAGAACAGCGCGCAAGTTGTCTCGGTGTTTGTGTGTCTAGAAATCTCGAGCTGCA 1432
Db 243 TCAGAACAGCGCGCAAGTTGTCTCGGTGTTTGTGTGTCTAGAAATCTCGAGCTGCA 302

QY 1433 GTAGCAACTCAACCTGACATCTGGCTCCAAAACAAAATCTCTTCTTGTGTATC 1492
Db 303 GTAGCAACTCAACCTGACATCTGGCTCCAAAACAAAATCTCTTCTTGTGTATC 362
QY 1493 TGACACGTCTGTGATGAATGTGAAAAAACATTAAGCTGCACAGACACCGTACTGCC 1552
Db 363 TGACACGTCTGTGATGAATGTGAAAAAACATTAAGCTGCACAGACACCGTACTGCC 422
QY 1553 AAGGAATTCATCTACCTCCAGGTGCCAGAGACATCTCCACCTGCTGTGAGCTGC 1612
Db 423 AAGGAATTCATCTACCTCCAGGTGCCAGAGACATCTCCACCTGCTGTGAGCTGC 482
QY 1613 ATGACTTCTCTGGAAGCTGTGTGTGCCAAGACAGCTGAGCTGTGTGTGCCAG 1672
Db 483 ATGACTTCTCTGGAAGCTGTGTGTGCCAAGACAGCTGAGCTGTGTGTGCCAG 542
QY 1673 CCAGAGCTGAGAGACATACACAGAGAACCTCTGCAACACAGCTTCACTACCTCG 1732
Db 543 CCAGAGCTGAGAGACATACACAGAGAACCTCTGCAACACAGCTTCACTACCTCG 602
QY 1733 TGGCCAGTCCATACCCAGCCAGACCTGTACTTGGCTCTTCTTCCCGGAGGCTCTA 1792
Db 603 TGGCCAGTCCATACCCAGCCAGACCTGTACTTGGCTCTTCTTCCCGGAGGCTCTA 662
QY 1793 TCAAGCAGTCCAGGTGAAGCAGAACATCTCGGTGACCTTGACACTTTGGCCCAAGT 1852
Db 663 TCAAGCAGTCCAGGTGAAGCAGAACATCTCGGTGACCTTGACACTTTGGCCCAAGT 722
QY 1853 TCACAACAGAGGCTTCAGAGCGGTCTGACGGTCTCTTATTAATTTCAAGAGG 1912
Db 723 TCACAACAGAGGCTTCAGAGCGGTCTGACGGTCTCTTATTAATTTCAAGAGG 782
QY 1913 AAGGCTTTTCACGGTGACCCCTGACACAAAAGAGTCTACTGAGAGACCCCACT 1972
Db 783 AAGGCTTTTCACGGTGACCCCTGACACAAAAGAGTCTACTGAGAGACCCCACT 842
QY 1973 GGGACCGGGGCTGTCATCCCTCACTCTGTCTCTGGAACATCAGCGTGCACAGACC 2032
Db 843 GGGACCGGGGCTGTCATCCCTCACTCTGTCTCTGGAACATCAGCGTGCACAGACC 902
QY 2033 AGGTGGCTGCTGACTTCTTTTAAAGAGCGAGCGGTGTCTGCGACAGAGGCGCG 2092
Db 903 AGGTGGCTGCTGACTTCTTTTAAAGAGCGAGCGGTGTCTGCGACAGAGGCGCG 962
QY 2093 CATTATGATCATCCAGAGCAGCGGACCCGGGCTGAGAGATCTTCAAGCTGACAGG 2152
Db 963 CATTATGATCATCCAGAGCAGCGGACCCGGGCTGAGAGATCTTCAAGCTGACAGG 1022
QY 2153 ATGTGCTCCCAAGCCAGACTTCCACATCAAGCTTGGGTGCAACATCTTAATGCA 2212
Db 1023 ATGTGCTCCCAAGCCAGACTTCCACATCAAGCTTGGGTGCAACATCTTAATGCA 1082
QY 2213 GCCCAGAGCGGCAAGAGAGTGAAGCTGCTCTTCTCGGTGACATTAACCCCAAGACTG 2272
Db 1083 GCCCAGAGCGGCAAGAGAGTGAAGCTGCTCTTCTCGGTGACATTAACCCCAAGACTG 1142
QY 2273 TGGACTTGACTGTATCTCTCATCGCAGCGGTGGAGTGTAACTTAACTGCTGTGCC 2332
Db 1143 TGGACTTGACTGTATCTCTCATCGCAGCGGTGGAGTGTAACTTAACTGCTGTGCC 1202
QY 2333 TCGGGCTCATATTTGCTGTGAAAAA 2360
Db 1203 TCGGGCTCATATTTGCTGTGAAAAA 1230
RESULT 13
BC021099 1410 bp mRNA linear PRI 14-MAY-2003
LOCUS BC021099
DEFINITION Homo sapiens CUB domain-containing protein 1, transcript variant 2,
mRNA (cDNA clone MGC:31813 IMAGE:459054), complete cds.
ACCESSION BC021099
VERSION BC021099.1 GI:18088978

KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
JOURNAL

REMARK
COMMENT

FEATURES
source

gene
/db_xref="LocusID:64866"
/db_xref="LocusID:64866"
/db_xref="LocusID:64866"

CDS

136..1167

/codon_start=1

/product="CUB domain-containing protein 1, isoform 2"

/protein_id="AAH21099.1"

/db_xref="GI:18088979"

/db_xref="LocusID:64866"

/translation="MAGLNCVSTALLVLLGAAKLRPRGAFAETALPRESNTYVLI
KLGPTLLAKPQYVIVISRHITMISIKGERIVTFPSCSPENHVFIEIQKIDWGS
PCPEEVLQPSSTSLPLNRTFIVDVAHKSIGLEQFSIPRLQIGRESCPDGV
HSISGRIDATVRIETPCSNQVSRIRKQEGVAMLHPMFIPRVSGFSIARSSIK
RLCIESVPEEGSGSTLMSANVPEGFPELMTQFVPAHRAVSGFSINFLNSGR
KEERREYVIRGSTINPEVFKLEDRKQPMNMGAFNLSLGGCDSDASPGILRQFQVLY
OHPQNESE"

BASE COUNT 321 a 395 c 370 g 324 t

ORIGIN

Query Match 18.8% Score 1159; DB 9; Length 1410;
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
Matches 1159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

148 GGGCGGGGCTCGGGCGGTCGCCCGCGCGAGTAGAGCCAGGCGGAGCGCAGCT 207
1 GGGCGGGGCTCGGGCGGTCGCCCGCGCGAGTAGAGCCAGGCGGAGCGCAGCT 60
208 GGGCGGGGCTCGGGCGGTCGCCCGCGCGAGTAGAGCCAGGCGGAGCGCAGCT 267
61 GGGCGGGGCTCGGGCGGTCGCCCGCGCGAGTAGAGCCAGGCGGAGCGCAGCT 120
268 AGGCGGTCGGGAGTATGAGCGGCGGCGGAGTATGAGCGGCGGAGTATGAGCGG 327
121 AGGCGGTCGGGAGTATGAGCGGCGGCGGAGTATGAGCGGCGGAGTATGAGCGG 180
328 CTGCTGCTGGGAGTATGAGCGGCGGCGGAGTATGAGCGGCGGAGTATGAGCGG 387
181 CTGCTGCTGGGAGTATGAGCGGCGGCGGAGTATGAGCGGCGGAGTATGAGCGG 240
388 CGAGGAAGCAATTACAGCTTCTAATAAGCTGGGAGCCCGGAGTATGAGCGG 447
241 CGAGGAAGCAATTACAGCTTCTAATAAGCTGGGAGCCCGGAGTATGAGCGG 300
448 TGTTCATGCTGATTTCTAATAAGCAATTACAGCTTCTAATAAGCTGGGAGG 507
301 TGTTCATGCTGATTTCTAATAAGCAATTACAGCTTCTAATAAGCTGGGAGG 360
508 ATAGCTTTACCTTTAGCTGAGGAGTCTCTGAGAACTCTTGTATAGATTCAGAA 567
361 ATAGCTTTACCTTTAGCTGAGGAGTCTCTGAGAACTCTTGTATAGATTCAGAA 420
568 AATATTGACTGATGATGAGGAGTCTCTTGGGAGGTTGAGTTCAGCTTCAGCA 627
421 AATATTGACTGATGATGAGGAGTCTCTTGGGAGGTTGAGTTCAGCTTCAGCA 480
628 TGTGTTGGCTTACCTTCAAGCAAGCACTTTCATCTGGGAGTTCAGGAGTTC 687
481 TGTGTTGGCTTACCTTCAAGCAAGCACTTTCATCTGGGAGTTCAGGAGTTC 540
688 GGTTTAGAGCTCAGTTTTCATCTCTGCTGAGGAGTTCAGGAGTTCAGGAGTTC 747
541 GGTTTAGAGCTCAGTTTTCATCTCTGCTGAGGAGTTCAGGAGTTCAGGAGTTC 600
748 CGAGGAGGATCATCTCATCTCATGAGGAGGAGTTCAGGAGTTCAGGAGTTC 807
601 CGAGGAGGATCATCTCATCTCATGAGGAGGAGTTCAGGAGTTCAGGAGTTC 660
808 ACCCTTCTGAGGAAGTTCAGGAGTTCAGGAGTTCAGGAGTTCAGGAGTTCAGG 867
661 ACCCTTCTGAGGAAGTTCAGGAGTTCAGGAGTTCAGGAGTTCAGGAGTTCAGG 720
868 TTACACCTCCCATGATTCACACCCAGAAATGCTCCGAGTTCAGGAGTTCAGG 927
721 TTACACCTCCCATGATTCACACCCAGAAATGCTCCGAGTTCAGGAGTTCAGG 780
928 TGTATTAAGAGTCTGATCATCTCATCTCATGAGGAGTTCAGGAGTTCAGGAG 987

```
Db      761 TCTATTAACGCTGTGATCATCATGCTGTGTTGAGGAGTGAAGGCTCAGCAACCCCTG 840
Qy      988 ATGTGTCACCACTACCCAGAAAGGCTTCCCTGAGAGATAGACTCATGACGCGGAGTTTGC 1047
Db      841 ATGTGTCACCACTACCCAGAAAGGCTTCCCTGAGAGATAGACTCATGACGCGGAGTTTGC 900
Qy      1048 GTTCTGTGACACCTGTGGGGCCAGCGTCTCTTCTTCACTTCAACCTTCCAACTGTGAG 1107
Db      901 GTTCTGTGACACCTGTGGGGCCAGCGTCTCTTCTTCACTTCAACCTTCCAACTGTGAG 960
Qy      1108 AGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1167
Db      961 AGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Qy      1168 AAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1227
Db      1021 AAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Qy      1228 TGTGACCAAGATGCTCCAAAGTCCAGGAGATCTCCGCTGACAGTTCCAAAGTTTGTCAA 1287
Db      1081 TGTGACCAAGATGCTCCAAAGTCCAGGAGATCTCCGCTGACAGTTTCCAAAGTTTGTCAA 1140
Qy      1288 CATCCACAAATGAAAGCA 1306
Db      1141 CATCCACAAATGAAAGCA 1159

RESULT 14
AX464028 1376 bp DNA linear PAT 16-JUL-2002
LOCUS AX464028
DEFINITION Sequence 161 from Patent WO0140466.
ACCESSION AX464028
VERSION AX464028.1 GI:21899031
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Baker,K.P., Beresini,M., Deforge,L., Desnoyers,L., Filvaroff,E.,
            Gao,W.O., Gerlstein,M.E., Goddard,A., Godowski,P.J., Gurney,A.L.,
            Sherwood,S., Smith,V., Stewart,T.A., Tumas,D., Watanabe,C.K.,
            Wood,W.L. and Zhang,Z.
            Secreted and transmembrane polypeptides and nucleic acids encoding
            same
            Patent: WO 0140466-A 161 07-JUN-2001;
JOURNAL
Genetech Inc. (US)
FEATURES
            source
            1..1376
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
BASE COUNT 311 a 384 c 352 g 329 t
ORIGIN
Query Match 18.1%; Score 1117; DB 6; Length 1376;
Best Local Similarity 100.0%; Fred.No. 0;
Matches 1117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      190 CCAGGCGGAGCGAGCTGCGCGGAGCTTGGGCGCTGCGGCGCGGCTCCACCGCTCG 249
Db      1 CCAAGGCGGAGCGAGCTGCGCGGAGCTTGGGCGCTGCGGCGCGGCTCCACCGCTCG 60
Qy      250 TTTTCCCAACCGAGCGGAGCGTCCCGAGATCATGCGCGGAGCTTGTCTT 309
Db      61 TTTTCCCAACCGAGCGGAGCGTCCCGAGATCATGCGCGGAGCTTGTCTT 120
Qy      310 ATCGACACTGCTAGGGGTTCTGCTGCTGGGTGGGCGCGGCTCCGCGGCGGAGCAAGCT 369
Db      121 ATCGACACTGCTAGGGGTTCTGCTGCTGGGTGGGCGCGGCTCCGCGGCGGAGCAAGCT 180
Qy      370 TTTGAGATTGCTCTGCGACGAGAAAGCAATTAAGTTCTCATAAAGCTGGGAGCCCG 429
```

```
Db      181 TTTGAGATTGCTCTGCGACGAGAAAGCAATTAAGTTCTCATAAAGCTGGGAGCCCG 240
Qy      430 ACTGTCGCGGAAACCCCTGTTACATGCTCAATTTCTTAAAGCATATACATGTTGCC 489
Db      241 ACTGTCGCGGAAACCCCTGTTACATGCTCAATTTCTTAAAGCATATACATGTTGCC 300
Qy      490 ATCAAGTCTGAGAAAGATAGTCTTTACCTTTAGCTCCAGAGTCTGAGATCACTTT 549
Db      301 ATCAAGTCTGAGAAAGATAGTCTTTACCTTTAGCTCCAGAGTCTGAGATCACTTT 360
Qy      550 GTCATAGAGATCCAGAAATATTTAGTGTATGTCAGGCCCATGTCCTTTTGGGAGTT 609
Db      361 GTCATAGAGATCCAGAAATATTTAGTGTATGTCAGGCCCATGTCCTTTTGGGAGTT 420
Qy      610 CAGCTTCAGCCCTGACATGCTGTGTGCTACCCCTCAACAGAACTTTTATCTGAGATGC 669
Db      421 CAGCTTCAGCCCTGACATGCTGTGTGCTACCCCTCAACAGAACTTTTATCTGAGATGC 480
Qy      670 AAAGCTCAATAGAGCATGCTTTAGAGCTGAGTTTTCATCCCTCGCTGAGAGAGATC 729
Db      481 AAAGCTCAATAGAGCATGCTTTAGAGCTGAGTTTTCATCCCTCGCTGAGAGAGATC 540
Qy      730 GGTCCGGGTGAGAGCTGCGCAGAGAGAGTCACTCTCATTCAGCGCCGAATGATGCC 789
Db      541 GGTCCGGGTGAGAGCTGCGCAGAGAGAGTCACTCTCATTCAGCGCCGAATGATGCC 600
Qy      790 ACCGTGTCAGAGATGCGAAACCTTCTGACAGCAATGGCACTGTGTCCCGATCAAGATGCA 849
Db      601 ACCGTGTCAGAGATGCGAAACCTTCTGACAGCAATGGCACTGTGTCCCGATCAAGATGCA 660
Qy      850 GAAGGAGTGAATAGGCTTACACCTCCCATAGTTCACCCAGAAATGTCTCCGGCTTC 909
Db      661 GAAGGAGTGAATAGGCTTACACCTCCCATAGTTCACCCAGAAATGTCTCCGGCTTC 720
Qy      910 AGCATTTGCAAAACCGCTCATATTAATAACGCTGTGTCATCATGAGTCTGTGTTAGAGGT 969
Db      721 AGCATTTGCAAAACCGCTCATATTAATAACGCTGTGTCATCATGAGTCTGTGTTAGAGGT 780
Qy      970 GAAAGCTCAGCAACCTGTATGTCTGCCAATCAACCAAGAGTTCCTCGAGATAGAGTTC 1029
Db      781 GAAAGCTCAGCAACCTGTATGTCTGCCAATCAACCAAGAGTTCCTCGAGATAGAGTTC 840
Qy      1030 ATGACGTGCGAGTTGTGTGTTCTGTCAGACCGCGGCGAGGTCCTCTCAACTTC 1089
Db      841 ATGACGTGCGAGTTGTGTGTTCTGTCAGACCGCGGCGAGGTCCTCTCTCAACTTC 900
Qy      1090 AACCTCTCCAACCTGTGAGAGAGAGAGAGCGGTTGAATCTACATCCCGGCTCCACC 1149
Db      901 AACCTCTCCAACCTGTGAGAGAGAGAGAGCGGTTGAATCTACATCCCGGCTCCACC 960
Qy      1150 ACCAACCCCGAGAGTGTTCAGCTGAGAGCAAGACGCTTGGAACATGCGCGGAACTTC 1209
Db      961 ACCAACCCCGAGAGTGTTCAGCTGAGAGCAAGACGCTTGGAACATGCGCGGAACTTC 1020
Qy      1210 AACCTCTCTGCAAGAGTGTGACCAAGATGCCAAATCTCCAGGATCTCTCGGCTGAG 1269
Db      1021 AACCTCTCTGCAAGAGTGTGACCAAGATGCCAAAGTCCAGGATCTCTCGGCTGAG 1080
Qy      1270 TTCCAAGTTTGTCTCAACATCCACAAATGAAAGCA 1306
Db      1081 TTCCAAGTTTGTCTCAACATCCACAAATGAAAGCA 1117

RESULT 15
AK026329 976 bp mRNA linear PRI 29-SEP-2000
LOCUS AK026329
DEFINITION Homo sapiens cDNA: FLJ22676 fis, clone HS110569.
ACCESSION AK026329
VERSION AK026329.1 GI:10439158
KEYWORDS AK026329.1
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
BASE COUNT
ORIGIN
Query Match
Best Local Similarity
Matches 959; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
14.7%; Score 909; DB 9; Length 976;
99.9%; Pred. No. 0;
1.976
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="HS110569"
/issue_type="human small intestine"
/clone_lib="HS1"
/note="Cloning vector pME18SFL3"
251 a 257 c 219 g 249 t
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
Ota,T., Suzuki,Y., Matsuura,A., Iwakura,S., Yamazaki,M., Tashiro,H.,
Nakamura,Y., Inoue,T., and Sugano,S.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 976)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Inoue,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Substitution
Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail:cdna1@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
Location/Qualifiers
1. 976
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="HS110569"
/issue_type="human small intestine"
/clone_lib="HS1"
/note="Cloning vector pME18SFL3"
251 a 257 c 219 g 249 t

Db
491
CTTAACTGAGGGCTCAAACTCTGAGCAAGTTGCTGGCTCTGAGACAGATTTTC
5496
TGAAGCTGCTGCTCAGTAAGAGGGCCAGCTGAGAACTGCTCTTTCTTTAAAGC
5555
541
TGAAGCTGCTGCTCAGTAAGAGGGCCAGCTGAGAACTGCTCTTTCTTTAAAGC
600
5556
CCAGGCCCACTTACATATAAATTTCAGGGTCACTGAGAAACAGTAAGTCCATTGTT
5615
601
CCAGGCCCACTTACATATAAATTTCAGGGTCACTGAGAAACAGTAAGTCCATTGTT
660
5616
GAGGCTTCTGATGAGCCAGCCCACTGCTATCCAGTGTCTGCTGCTGAGGAAG
5675
661
GAGGCTTCTGATGAGCCAGCCCACTGCTATCCAGTGTCTGCTGCTGAGGAAG
720
5676
GCCAGCATGAGAGCTGCTCTTAATGCTGTGCTCACTTGCACAGAAAGAGTCTC
5735
721
GCCAGCATGAGAGCTGCTCTTAATGCTGTGCTCACTTGCACAGAAAGAGTCTC
780
5736
AAGAAAGTCAACTGGAGCAAGCAAGCCCACTGAGCTTGTGTAAGGTTAGC
5795
781
AAGAAAGTCAACTGGAGCAAGCAAGCCCACTGAGCTTGTGTAAGGTTAGC
840
5796
AGACTGTGTGTGTGATCTGAGTCTTCACTGAAATTAATTTTCACTTGCAGATCT
5855
841
AGACTGTGTGTGTGATCTGAGTCTTCACTGAAATTAATTTTCACTTGCAGATCT
900
5856
TTTTAGTGGCAATTTATTTCAATTTCTGCTTTTAAATTAACAAATGTAACAAACAA
5915
901
TTTTAGTGGCAATTTATTTCAATTTCTGCTTTTAAATTAACAAATGTAACAAACAA
960

Search completed: February 20, 2004, 10:36:22
Job time : 21710 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: February 20, 2004, 03:53:40 ; Search time 11576 Seconds
(without alignments)
12939.572 Million cell updates/sec

Title: US-09-899-569A-3

Perfect score: 6163
Sequence: 1 ccaacgcgcgaatggggagc.....agctcgcacaaaaaa 6163

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 22781392 seqs, 1215238056 residues

Word size: 0

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

EST:
1: em_eetba:*
2: em_eetbam:*
3: em_eetbu:*
4: em_eetbu:*
5: em_eetbu:*
6: em_eetbu:*
7: em_eetbu:*
8: em_eetbu:*
9: em_eetbu:*
10: em_eetbu:*
11: em_eetbu:*
12: em_eetbu:*
13: em_eetbu:*
14: em_eetbu:*
15: em_eetbu:*
16: em_eetbu:*
17: em_eetbu:*
18: em_eetbu:*
19: em_eetbu:*
20: em_eetbu:*
21: em_eetbu:*
22: em_eetbu:*
23: em_eetbu:*
24: em_eetbu:*
25: em_eetbu:*
26: em_eetbu:*
27: em_eetbu:*
28: em_eetbu:*
29: em_eetbu:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	826	13.4	9 AB033417	AB033417 AB033417
2	826	13.4	9 AB063026	AB063026 EST392 hu
3	766	12.4	13 B0951236	B0951236 AGENCOURT
4	735	11.9	13 B0962997	B0962997 AGENCOURT

5	728	11.8	1008	13 B0691792	B0691792 AGENCOURT
6	720	11.7	868	13 B0681577	B0681577 AGENCOURT
7	706	11.5	882	13 B0153170	B0153170 AGENCOURT
8	704	11.4	1150	10 BE877551	BE877551 601486503
9	662	10.7	676	13 B0157152	B0157152 AGENCOURT
10	642	10.4	643	12 B0044552	B0044552 603622093
11	635	10.3	738	14 CA307237	CA307237 UI-H-F11-
12	629	10.2	851	10 B0419603	B0419603 602451816
13	628	10.2	683	12 B1457122	B1457122 603185367
14	621	10.1	712	12 B0018619	B0018619 UI-H-DH1-
15	612	9.9	871	13 B0889198	B0889198 AGENCOURT
16	611	9.9	611	10 B0403590	B0403590 602419303
17	594	9.6	717	14 CD367582	CD367582 UI-H-F11-
18	586	9.5	671	9 AL710125	AL710125 DKFZP686A
19	573	9.3	573	12 B0966953	B0966953 1128607.Y
20	570	9.2	570	12 B0966788	B0966788 1128609.Y
21	565	9.2	565	12 B0893567	B0893567 1126009.X
22	562	9.1	562	14 CB157642	CB157642 K-EST0216
23	559	9.1	690	14 CB128689	CB128689 K-EST0178
24	553	9.0	977	13 B0950224	B0950224 AGENCOURT
25	546	8.9	552	9 AW138207	AW138207 UI-H-B11-
26	545	8.8	863	10 BE877911	BE877911 601489392
27	537	8.7	614	14 CB122885	CB122885 K-EST0171
28	537	8.7	692	9 AL601433	AL601433 DKFZP313H
29	535	8.7	723	14 CA314193	CA314193 UI-CF-FNO
30	530	8.6	530	12 B0840003	B0840003 K-EST0116
31	522	8.5	522	12 B0893690	B0893690 1128607.X
32	514	8.3	916	13 B0691185	B0691185 AGENCOURT
33	513	8.3	513	14 CB129853	CB129853 K-EST0179
34	513	8.3	843	10 B0499230	B0499230 602545670
35	505	8.2	505	12 B0840669	B0840669 K-EST0117
36	504	8.2	504	12 B0018832	B0018832 UI-H-DH1-
37	503	8.2	621	10 B0119833	B0119833 602352174
38	495	8.0	495	12 B0701518	B0701518 UI-E-E10-
39	488	7.9	594	9 AW578426	AW578426 RC1-CT024
40	488	7.9	594	9 AW604363	AW604363 RC1-CT024
41	486	7.9	542	10 BF336602	BF336602 RC1-CT024
42	480	7.8	480	13 BX473477	BX473477 DKFZP686C
43	479	7.8	481	9 A1275971	A1275971 QW06G08.X
44	478	7.8	478	12 B0838647	B0838647 K-EST0115
45	476	7.7	527	12 B0826981	B0826981 K-EST0099

ALIGNMENTS

RESULT 1
AB033417
LOCUS
DEFINITION
AB033417 Homo sapiens skov Homo sapiens cDNA, mRNA sequence.
VERSION
AB033417.1 GI:6062545
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Unpublished
Contact: Zhou Yong
Department of Reproductive Physiology and Endocrinology
Medical Institute of Bioregulation, Kyushu University
Tsurumihara 4546, Beppu, Oita 874-0838, Japan
Tel: 0977-27-1660
Fax: 0977-27-1661
Email: yzhou@surumi.beppu.kyushu-u.ac.jp.

FEATURES
source
1..933
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Db	481	TTGATCCATCCAGACGGCTTTGCTGAGTTCTAATAAGACGGGCATCTGAGACACCCGTCOG	540									
Qy	2855	TGTTCTTACCAAGAAATCTTAAGAAGAGAAATTAACAGAAAGAAACAGCAGAGAGTTT	2914									
Db	541	TGTTCTTACCAAGAAATCTTAAGAAGAGAAATTAACAGAAAGAAACAGCAGAGAGTTT	600									
Qy	2915	CCTGACACCCGCAACTTCCATTTGCTCAGTGGACTCAATTCCTAAGGCAAGCATTTGAA	2974									
Db	601	CCTGACACCCGCAACTTCCATTTGCTCAGTGGACTCAATTCCTAAGGCAAGCATTTGAA	660									
Qy	2975	ATGATGAATTTCCAAATCTGGATACAGTATGACAGCTCATGTGTCCTCAACTTAGGCTGT	3034									
Db	661	ATGATGAATTTCCAAATCTGGATACAGTATGACAGCTCATGTGTCCTCAACTTAGGCTGT	720									
Qy	3035	GCGGTTAGCCACGCTTGTAATGAGAGAGAGAGGCGCTGAGTCACTAGACATAGGTTTGAG	3094									
Db	721	GCGGTTAGCCACGCTTGTAATGAGAGAGAGAGGCGCTGAGTCACTAGACATAGGTTTGAG	780									
Qy	3095	CAAGCCCTGATTCAGAGTGTAAACAGAGGCTTCCCTCTTCAGACAAACAGTTCCAT	3154									
Db	781	CGAGCCCTGATTCAGAGTGTAAACAGAGGCTTCCCTCTTCAGACAAACAGTTCCAT	840									
Qy	3155	TCCAAGAGCCTACCTGAGAGTCCCTACTCTCACTGGG	3191									
Db	841	TCCAAGAGCCTACCTGAGAGTCCCTACTCTCACTGGG	877									
RESULT 3												
LOCUS	BO951236	984 bp	mRNA	linear	EST 21-AUG-2007							
DEFINITION	AGENCOURT_8817950 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6422238											
ACCESSION	BO951236											
VERSION	BO951236.1	GI:22366714										
KEYWORDS	EST.											
SOURCE	Homo sapiens (human)											
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.											
REFERENCE	1 (bases 1 to 984)											
AUTHORS	NIH-MGC http://mgi.nci.nih.gov/.											
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)											
JOURNAL	Unpublished											
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: DCTP/NIH/Gatzert cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov Plate: L1CM2558 row: 0 column: 07 High quality sequence start: 6 High quality sequence stop: 719. Location/Qualifiers 1. 984 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6422238" /tissue_type="large cell carcinoma" /lab_host="DH10B (phage-resistant)" /clone_11b="NIH MGC 18" /note="Organ: lung; Vector: pORF7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGACG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."											
FEATURES	Source											
BASE COUNT	238 a	301 c	247 g	197 t	1 others							

ORIGIN	Query Match	12.4%	Score 766	DB 13	Length 984
Best Local Similarity	100.0%	Pred. No. 4,6e-156			
Matches 766	Conservative 0	Mismatches 0	Indels 0	Gaps 0	
QY	CTGTGTTTGGAGGCTGGAAGGCTCCAGCAACCTTGATGCTCCCACTACCCAGGAAGGCTTCC				1015
DB	23 CTGTGTTTGGAGGCTGGAAGGCTCCAGCAACCTTGATGCTCCCACTACCCAGGAAGGCTTCC				82
QY	CTGAGATGAGAGCTCATGACGCGGAGGATTTGTCGTTCTTGACACCTCAGGAGCAAGCTCT				1075
DB	83 CTGAGATGAGAGCTCATGACGCGGAGGATTTGTCGTTCTTGACACCTCAGGAGCAAGCTCT				142
QY	1076 CTTTCTTCAACTTCAACCTCTCCACTGTGAGAGGAAGAGAGCGGGTTGAATACTACA				1135
DB	143 CTTTCTTCAACTTCAACCTCTCCACTGTGAGAGGAAGAGAGCGGGTTGAATACTACA				202
QY	1136 TCCCGGGGCTCCACCAACCCCGAGGTGTTTCAAGCTGAGAGACACAGCTCTGGAAACA				1195
DB	203 TCCCGGGGCTCCACCAACCCCGAGGTGTTTCAAGCTGAGAGACACAGCTCTGGAAACA				262
QY	1196 TGGCGGGGAACTTCAACCTCTCTGAGAGGCTGTGACCAAGATGCCCAAGTCCAGGGA				1255
DB	263 TGGCGGGGAACTTCAACCTCTCTGAGAGGCTGTGACCAAGATGCCCAAGTCCAGGGA				322
QY	1256 TCTTCGCGCTGCAAGTTTGGTTCACATCCACAAAATGAAAGCAATAAAATCT				1315
DB	323 TCTTCGCGCTGCAAGTTTGGTTCACATCCACAAAATGAAAGCAATAAAATCT				382
QY	1316 ACGTGTGTTGACTGATGATGAGGAGGACATGTCTACCATCGAGCCACCGCCCTCA				1375
DB	383 ACGTGTGTTGACTGATGATGAGGAGGACATGTCTACCATCGAGCCACCGCCCTCA				442
QY	1376 AACAGAGCGCGCAAGTTTGTCCCTGCGGCTTTGCGTGCTAGAGATCGCGACCTGCAGTA				1435
DB	443 AACAGAGCGCGCAAGTTTGTCCCTGCGGCTTTGCGTGCTAGAGATCGCGACCTGCAGTA				502
QY	1436 GCAACCTCACCTGACATCTGGCTCCAAACAACAATACTCTCTTGTGATGATCTGA				1495
DB	503 GCAACCTCACCTGACATCTGGCTCCAAACAACAATACTCTCTTGTGATGATCTGA				562
QY	1496 CACGTCGTGTGATGAATGTGAAAAAACAATAAGCTGCACAGACACACCGGTACTGCCAAA				1555
DB	563 CACGTCGTGTGATGAATGTGAAAAAACAATAAGCTGCACAGACACACCGGTACTGCCAAA				622
QY	1556 GGAATTCCTACTCACTCCAGGTGCCCGAGTCAATCTCCACCTGCGCTGTGAGCTGATG				1615
DB	623 GGAATTCCTACTCACTCCAGGTGCCCGAGTCAATCTCCACCTGCGCTGTGAGCTGATG				682
QY	1616 ACTTCTCTGGAAGCTCTGGTGCCCAAGGACAGAGCTCAGCTGTGCTGTGTGACAGCC				1675
DB	683 ACTTCTCTGGAAGCTCTGGTGCCCAAGGACAGAGCTCAGCTGTGCTGTGTGACAGCC				742
QY	1676 AGAAGCTGCAGACAGCATACACAGAGAAAGCCCTTGCAACACACAGCTT				1721
DB	743 AGAAGCTGCAGACAGCATACACAGAGAAAGCCCTTGCAACACACAGCTT				788
RESULT 4					
LOCUS	BQ962997	928 bp	mRNA	linear	EST 21-AUG-2002
DEFINITION	AGNCSCOURT 8778970 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:6378668				
ACCESSION	BQ962997				
VERSION	BQ962997.1	GI:22378463			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
REFERENCE	Mammalia; Euteleostomi; Primates; Catarrhini; Homiidae; Homo.				
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .				

TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished
COMMENT	Contact: Robert Strausberg, Ph.D. National Cancer Institute

5289 CTGACATTTCCCAACCTCCAGCCCTTCCAGAGAGAGTATGAGTCCCTGATTCAC 5348
69 CTGACATTTCCCAACCTCCAGCCCTTCCAGAGAGAGTATGAGTCCCTGATTCAC 128
5349 CAGGTGGATTTGGCTTCTTGGCTGGCTATCTGTCACTCAACCATCACTGTTG 5408
129 CAGGTGGATTTGGCTTCTTGGCTGGCTATCTGTCACTCAACCATCACTGTTG 188
5409 CCTGAGAGAGACCAAGTGGGATTTCTTCACTGAGGGGCTCAAACTCCTGAGCAAG 5468
189 CCTGAGAGAGACCAAGTGGGATTTCTTCACTGAGGGGCTCAAACTCCTGAGCAAG 248
5469 TTGCTGGCTCTGAGACCAAGTATTTCTGAGAGCTGTGCTCAGTAAAGGGGCCAGCTG 5528
249 TTGCTGGCTCTGAGACCAAGTATTTCTGAGAGCTGTGCTCAGTAAAGGGGCCAGCTG 308
5529 AGGAACCTGGCTCTTCTTTTAAAGCCAGGCCCACTTACATTAACATTTAGGGTC 5588
309 AGGAACCTGGCTCTTCTTTTAAAGCCAGGCCCACTTACATTAACATTTAGGGTC 368
5589 ACTGAAACAGTAAAGTGGCTATTTGTGAAGCTATCTGATGCCAGCCCACTGCTATCC 5648
369 ACTGAAACAGTAAAGTGGCTATTTGTGAAGCTATCTGATGCCAGCCCACTGCTATCC 428
5649 AGTGGTCTGACCTGCTTACAGAGAGAGCCAGGATGAGGAGTCTCTATATGCTGT 5708
429 AGTGGTCTGACCTGCTTACAGAGAGAGCCAGGATGAGGAGTCTCTATATGCTGT 488
5709 GGTCTATTCACAGAGAGAGAGAGTCTCAAGAGAGTCAATCTGAGACCAAGAGCCCA 5768
489 GGTCTATTCACAGAGAGAGAGAGTCTCAAGAGAGTCAATCTGAGACCAAGAGCCCA 548
5769 CCGGACATAGGCTTGTAAAGTGTAGCAGAGTGTGTGTGTGTGTGTGTGTGTGTGT 5828
549 CCGGACATAGGCTTGTAAAGTGTAGCAGAGTGTGTGTGTGTGTGTGTGTGTGTGT 608
5829 GGAATTAATTTATTTATTCATTCAGATGATCTTTTATGAGGATTTTATTCATTTCTG 5888
609 GGAATTAATTTATTTATTCATTCAGATGATCTTTTATGAGGATTTTATTCATTTCTG 668
5889 TAAATTAACCAATGTATCCAAAAACAAGTATCAAGCTGTTTAAAGTCTGCTGCTAT 5948
669 TAAATTAACCAATGTATCCAAAAACAAGTATCAAGCTGTTTAAAGTCTGCTGCTAT 728
5949 CCGCTGTTCACTAAGAGCCCGGTTTCCAGTGTGTGATCTGTGACAGGCTGAGCATGG 6008
729 CCGCTGTTCACTAAGAGCCCGGTTTCCAGTGTGTGATCTGTGACAGGCTGAGCATGG 788
6009 CTCAGCAGATGCTGCTTAAATTTGTGATGATACAGAAAGCCAGGCTTTGGATTCAG 6067
789 CTCAGCAGATGCTGCTTAAATTTGTGATGATACAGAAAGCCAGGCTTTGGATTCAG 847

RESULT 6
BO687577 868 bp mRNA linear EST 15-JUL-2002
LOCUS BO687577
DEFINITION AGENCOURT 8343949 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6248364
5', mRNA sequence.
ACCESSION BO687577
VERSION BO687577.1 GI:21812893
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 868)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
CONTACT: Robert Strausberg, Ph.D.
COMMENT Email: cgaabs-r@mail.nih.gov
Tissue Procurement: ATCC

cdna Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLI)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNLI at:
http://image.llnl.gov
Plate: L1CM2388 row: b column: 13
High quality sequence stop: 702.
Location/Qualifiers
1. 868
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6248364"
/tissue_type="ductal carcinoma, cell line"
/lab_host="pDH10B (phage-resistant)"
/clone_id="NIH_MGC_110"
/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

BASE COUNT 214 a 238 c 206 g 210 t

Query Match 11.7%; Score 720; DB 13; Length 868;
Best Local Similarity 100.0%; Pred. No. 4.1e-16;
Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

5065 CAAGACCTCTGGAACATCTTCCAGAGATAGATTTTAAGCATGCCCTGGCGCTGGC 5124
1 CAAGACCTCTGGAACATCTTCCAGAGATAGATTTTAAGCATGCCCTGGCGCTGGC 60
5125 TATGACACCTTCCCTTCTGAAAGCTGTCTCTGACAGTGAACCTTGAGC 5184
61 TATGACACCTTCCCTTCTGAAAGCTGTCTCTGACAGTGAACCTTGAGC 120
5185 CGAGATGCTGACCTGATTAAGAGGCGCAAGAGAGGCTGCGCTTCTTCCCTCACTGA 5244
121 CGAGATGCTGACCTGATTAAGAGGCGCAAGAGAGGCTGCGCTTCTTCCCTCACTGA 180
5245 GAGCCTTATTTGAATTCATCTGTGTGAGGAGCCCTTCCATTTTGTGACATTTCCCAAC 5304
181 GAGCCTTATTTGAATTCATCTGTGTGAGGAGCCCTTCCATTTTGTGACATTTCCCAAC 240
5305 CTCCAGGCGCTTCCAGACAGAGTAAAGTGCCTGATTTCCACCCAGAGTGGATTTGGCC 5364
241 CTCCAGGCGCTTCCAGACAGAGTAAAGTGCCTGATTTCCACCCAGAGTGGATTTGGCC 300
5365 TTCTTAAGGCTGCTACTTGTTCACCATCAAGATCACTGTGCTGCAAGACACCAAC 5424
301 TTCTTAAGGCTGCTACTTGTTCACCATCAAGATCACTGTGCTGCAAGACACCAAC 360
5425 GTGGCATTCTTCTTCAACTGAGGCTCAAAACTCTGAGCAAGTGTGCTGCTGAGA 5484
361 GTGGCATTCTTCTTCAACTGAGGCTCAAAACTCTGAGCAAGTGTGCTGCTGAGA 420
5485 CCAGATTTCTGAGAGCTGTGCTGAGTGAAGGGGCCAGGCTGAGAAACCTGGCTCTT 5544
421 CCAGATTTCTGAGAGCTGTGCTGAGTGAAGGGGCCAGGCTGAGAAACCTGGCTCTT 480
5545 TTCTTAAGGCGCAGGCCCACTTACATTAACATTTTCAAGGCTCACTGAGAAACAGTGAAG 5604
481 TTCTTAAGGCGCAGGCCCACTTACATTAACATTTTCAAGGCTCACTGAGAAACAGTGAAG 540
5605 TGCCATTTGTGGAAGCTTACATGATGAGCCCACTGCTCATCCAGTGTGCTGCAATGC 5664
541 TGCCATTTGTGGAAGCTTACATGATGAGCCCACTGCTCATCCAGTGTGCTGCAATGC 600
5665 CTACAGAGAGGCGCAGGATGAGGATGAGTGTCTAATGTGTGTGATTCAGACAGAG 5724

DB 601 CTACGAGAAAGCCACCGATGAGACTGGTCTTAATGCTGTGTCTCAATTGACAGAG 660
QY 5725 GGAAAGTCTCAAGAAAGTCAACTGGAGCAAGCAAGCCACCGGACATGGCTTGG 5784
DB 661 GGAAAGTCTCAAGAAAGTCAACTGGAGCAAGCAAGCCACCGGACATGGCTTGG 720
RESULT 7
BUI53170 882 bp mRNA linear EST 03-SEP-2002
LOCUS AGENCOURT 7937392 NIH_MGC_70 Homo sapiens cDNA IMAGE:6026162
DEFINITION 5' mRNA sequence.
ACCESSION BUI53170 GI:22666702
VERSION BUI53170.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS 1 (bases 1 to 882)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL3239 row: h column: 03
High quality sequence stop: 645.
Location/Qualifiers
1..882
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6026162"
/issue_type="epithelioid carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_70"
/note="Organ: Pancreas; Vector: pCMV-Sport6; Site: 1: NCI;
Site 2: Sall; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.1 kb. Library constructed by Life
Technologies."

BASE COUNT 205 a 258 c 184 g 235 t
ORIGIN
Query Match 11.5%; Score 706; DB 13; Length 882;
Best Local Similarity 99.8%; Pred. No. 4.3e-143;
Matches 806; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4630 AATAGCTGGATCACTGGCAAAACCAACCAATGCGCCAGCTAATTTGTATTTTGTAGAG 4689
DB 3 AATAGCTGGATCACTGGCAAAACCAACCAATGCGCCAGCTAATTTGTATTTTGTAGAG 62
QY 4690 AAGAGGCTTACACATGTCGCCAGGCTGTCTCAACCTCTGGGCTCAAGCAATCTCTCT 4749
DB 63 AAGAGGCTTACACATGTCGCCAGGCTGTCTCAACCTCTGGGCTCAAGCAATCTCTCT 122
QY 4750 GCGTGGCTCCCAAAAGTCTGGAGTTAAGATGTAGCAACCGCATCCAGCCCAAC 4809
DB 123 GCGTGGCTCCCAAAAGTCTGGAGTTAAGATGTAGCAACCGCATCCAGCCCAAC 182
QY 4810 CTCATTATACCAATTAACCTGCGCACTGAAGCTTTTGTCTTCAACCTGCTCT 4869
DB 183 CTCATTATACCAATTAACCTGCGCACTGAAGCTTTTGTCTTCAACCTGCTCT 242
QY 4870 GATCTGGAAGAGAGGATTTATTTACTTGTACGACAGTCCCAAGTCAATTTTC 4929

DB 243 GATCTGGAAGAGAGGATTTATTTATAGCTTTGACGACAGTCCCAAGTCAATTTTC 302
QY 4930 TCGGCAAAAACCTCTTCAAAAATAATGACTTCAATTGATTAATGATCACTT 4989
DB 303 TCGGCAAAAACCTCTTCAAAAATAATGACTTCAATTGATTAATGATCACTT 362
QY 4990 GGAAATGACCGGCTCACTTGTTCATAGCATTAAGAAAGAAATTTATAGTCTCT 5049
DB 363 GGAAATGACCGGCTCACTTGTTCATAGCATTAAGAAAGAAATTTATAGTCTCT 422
QY 5050 AAATGGCTGTACTGACAGACCTCTTGAACACTTCCAGAGATAGATTAATTAATGAT 5109
DB 423 AAATGGCTGTACTGACAGACCTCTTGAACACTTCCAGAGATAGATTAATGAT 482
QY 5110 GCCCTTGGGCTTGCCTATGACACCTTTCCTTGAAGATCTGTCTCTGCCAGTACC 5169
DB 483 GCCCTTGGGCTTGCCTATGACACCTTTCCTTGAAGATCTGTCTCTGCCAGTACC 542
QY 5170 CTTGGCTTTGAGAGCGGAGATGCTGACCCCTGATTAAGGGCCAAAGAGGCTGGGCTT 5229
DB 543 CTTGGCTTTGAGAGCGGAGATGCTGACCCCTGATTAAGGGCCAAAGAGGCTGGGCTT 602
QY 5230 CTTGCCCTCACTGAGAGGCTTATTTGAATTCATCTGTGAGAGCTTACCTCATTC 5289
DB 603 CTTGCCCTCACTGAGAGGCTTATTTGAATTCATCTGTGAGAGCTTACCTCATTC 662
QY 5290 TCGACATTTCCCAACCTCCAGCCCTTCAAGAGAGCTAGTGCCTGACATTCACCC 5349
DB 663 TCGACATTTCCCAACCTCCAGCCCTTCAAGAGAGCTAGTGCCTGACATTCACCC 722
QY 5350 AAGGTGGATTTGGCTTCTTCTTATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5409
DB 723 AAGGTGGATTTGGCTTCTTCTTATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 782
QY 5410 CTGCAAGACACCAAGTGGCAATTTTC 5437
DB 783 CTGCAAGACACCAAGTGGCAATTTTC 810
RESULT 8
BE877551 1150 bp mRNA linear EST 20-OCT-2000
LOCUS BE877551
DEFINITION 601486503P1 NIH_MGC_69 Homo sapiens cDNA clone IMAGE:3888768 5',
mRNA sequence.
ACCESSION BE877551
VERSION BE877551.1 GI:10326327
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS 1 (bases 1 to 1150)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP/Gazdar
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL9669 row: f column: 01
High quality sequence stop: 724.
Location/Qualifiers
1..1150
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3888768"
/issue_type="Large cell carcinoma, undifferentiated"

```

/lab host="DH10B (phage-resistant)"
/clone_id="NH_MGC_69"
/notes="Organ: lung; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.1 kb. library constructed by Life
Technologies."
BASE COUNT      259 a      329 c      312 g      250 t
ORIGIN

```

Query Match	11.4%	Score 704	DB 10	Length 1150
Best Local Similarity	100.0%	Pred. No. 9.5e-143		
Matches 704	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	2412	CATCAATATCTGAGATGCCGAGGACGCCAAAAAGTTTCAGAAAGGCCGAAAGACAATGA	2471	
Db	1	CATCAATATCTGAGATGCCGAGGACGCCAAAAAGTTTCAGAAAGGCCGAAAGACAATGA	60	
QY	2472	CTCCCATGTGTATGTGACGTCAATCGAGGACCAATGTATATGGGCATCTGCTACAGAGATTC	2531	
Db	61	CTCCCATGTGTATGTGACGTCAATCGAGGACCAATGTATATGGGCATCTGCTACAGAGATTC	120	
QY	2532	CAGGGGCTCCCTCCGCGACGACAGAGGTGACACCTACCCGGCGTTCCAGGGGACACCATGG	2591	
Db	121	CAGGGGCTCCCTCCGCGACGACAGAGGTGACACCTACCCGGCGTTCCAGGGGACACCATGG	180	
QY	2592	GGTGTGTCTTCCTCCCTCCACCACCACCATATGTGCTCCAGGGGCCCAACTGCAAAAGTTGGCCAC	2651	
Db	181	GGTGTGTCTTCCTCCCTCCACCACCACCATATGTGCTCCAGGGGCCCAACTGCAAAAGTTGGCCAC	240	
QY	2652	TGAGAGGCCACCTCTCTGCTCCCTCCCTGAGTCTGAGGTGAAGTGAACCGTACACCTTTCTCCA	2711	
Db	241	TGAGAGGCCACCTCTCTGCTCCCTCCCTGAGTCTGAGGTGAAGTGAACCGTACACCTTTCTCCA	300	
QY	2712	TCCCAACCAATGGGAGTGTAAACAGACAGAAACAACAGACTTTCCTTAACCTGAACACTCAGGA	2771	
Db	301	TCCCAACCAATGGGAGTGTAAACAGACAGAAACAACAGACTTTCCTTAACCTGAACACTCAGGA	360	
QY	2772	GCCCATGAGGACGACAGAGATTAATCTGATCACTTCAGACGCTTGGCTGAGTTTCATAAG	2831	
Db	361	GCCCATGAGGACGACAGAGATTAATCTGATCACTTCAGACGCTTGGCTGAGTTTCATAAG	420	
QY	2832	CAGGGCACTGAGACACCCGCTCGTGTCTTAACAGAAATCCTAAAGAAAGAGAAATTATA	2891	
Db	421	CAGGGCACTGAGACACCCGCTCGTGTCTTAACAGAAATCCTAAAGAAAGAGAAATTATA	480	
QY	2892	CAGAAAGGAACGACGAGAGGTTTTCTCTGACACCCGCAACTTCACATTGCTCTAGTGAATC	2951	
Db	481	CAGAAAGGAACGACGAGAGGTTTTCTCTGACACCCGCAACTTCACATTGCTCTAGTGAATC	540	
QY	2952	ATTCTAAGGGGCAAGACATTTGAATAATGAAATTCCTAATCTGATACATTCATGACAGCTC	3011	
Db	541	ATTCTAAGGGGCAAGACATTTGAATAATGAAATTCCTAATCTGATACATTCATGACAGCTC	600	
QY	3012	ATGTGTCCTCTCAACTTAGGCTGTGGGGTTTACACAGCTGTATGAGAGAGAGAGGCTTG	3071	
Db	601	ATGTGTCCTCTCAACTTAGGCTGTGGGGTTTACACAGCTGTATGAGAGAGAGAGGCTTG	660	
QY	3072	AGTCACCTAGCATAGGAGTTTGACGACAGCCCTGATTCAGAGTGT	3115	
Db	661	AGTCACCTAGCATAGGAGTTTGACGACAGCCCTGATTCAGAGTGT	704	
RESULT 9				
LOCUS	BU157152	676 bp	mRNA	linear
DEFINITION	AGNCCOURT 7941310 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:601551			
ACCESSION	BU157152			
VERSION	BU157152.1	GI:22670684		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	(baase 1 to 676)	NIH-MGC http://mgc.nci.nih.gov/ .	National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished
	Contact: Robert Strausberg, Ph.D.			

Mammalia: *Utheria*: Primates; *Catarrhini*; *Hominidae*: *Homo*.
1 (baaes 1 to 676)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DTF/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLML13201 row: 9 column: 08
High quality sequence stop: 643.

FEATURES	Location/Qualifiers
source	1. .676

Query Match	Best Local Similarity	100.0%;	Score 662;	DB 13;	Length 676;
Matches 662;	Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;
BASE COUNT	170 a	180 c	158 g	168 t	
ORIGIN					
Query Match	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:6011511"				
	/issue_type="large cell carcinoma"				
	/lab_host="DH10B (phage-resistant)"				
	/clone_1b="NH_MGC_68"				
	/note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;				
	Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.				
	Average insert size 1.8 kb. Library constructed by Life				
	Technologies."				
QY	5251	TTATTGAATTCACGTGTGAGAGCCCTTACATTCCTGACATTTCCCACTTCCCA	5310		
Db	1	TTATTGAATTCACGTGTGAGAGCCCTTACATTCCTGACATTTCCCACTTCCCA	60		
QY	5311	GCCCTTCGAAGAGACTAGTGCCTCGATTCACCCAAAGGAGATTTGGCTTCTT	5370		
Db	61	GCCCTTCGAAGAGACTAGTGCCTCGATTCACCCAAAGGAGATTTGGCTTCTT	120		
QY	5371	AGGTGCTACTTGTGCATCAACCGACATCACTGTTGCTGCAGAGACACAGTGGCC	5430		
Db	121	AGGTGCTACTTGTGCATCAACCGACATCACTGTTGCTGCAGAGACACAGTGGCC	180		
QY	5431	ATTTTCTTCAACTGAGGGCTCAAAATCTCTGCAGCAAGTTGTGCTCTGAGACCACTA	5490		
Db	181	ATTTTCTTCAACTGAGGGCTCAAAATCTCTGCAGCAAGTTGTGCTCTGAGACCACTA	240		
QY	5491	TTTTCCTGAGACTGTGCTCAGTGAAGGGGCCAGCCCTGAGGAACCTGGCTCTTTCTT	5550		
Db	241	TTTTCCTGAGACTGTGCTCAGTGAAGGGGCCAGCCCTGAGGAACCTGGCTCTTTCTT	300		
QY	5551	AAAGCCAGGCCCCCACTTACATTAATAAATTTTCAGGGTCACTGGAACAAGTAAATGTCAT	5610		
Db	301	AAAGCCAGGCCCCCACTTACATTAATAAATTTTCAGGGTCACTGGAACAAGTAAATGTCAT	360		
QY	5611	TTGTTGAAGCTTACTGATGCAAGCCCACTGCTCATCAAGTGTCTGCCATGCTTACGA	5670		
Db	361	TTGTTGAAGCTTACTGATGCAAGCCCACTGCTCATCAAGTGTCTGCCATGCTTACGA	420		
QY	5671	GGAAGGCAAGGCACTGAGAGCTGGTCTCTAATCTGTGGTCAATTGCAAGAAAGGAAG	5730		
Db	421	GGAAGGCAAGGCACTGAGAGCTGGTCTCTAATCTGTGGTCAATTGCAAGAAAGGAAG	480		
QY	5731	GCTCAAGAAAGATCAATCTGGGACAAGCAACAAGCCACCGGACATGGCTTGGTAAAG	5790		
Db	481	GCTCAAGAAAGATCAATCTGGGACAAGCAACAAGCCACCGGACATGGCTTGGTAAAG	540		
QY	5791	TTAGCAAGCTGTGTGTGTGATCTGCAGTGTCTTCACTGGAATAATTTATTCATTGCA	5850		

```

Db      541 TTAGCAGACTGCTGTGTGTGATCTGACATGCTTCACTGAAATATTTATTCATTGCG 600
Oy      5851 ATACTTTTAGGTGATTTTATTCATTTCTGTCGCTTAAATTAACAATGTACCAAA 5910
Db      601 ATACTTTTAGGTGATTTTATTCATTTCTGTCGCTTAAATTAACAATGTACCAAA 660
Oy      5911 AA 5912
Db      661 AA 662

RESULT 10
BM044552      643 bp      mRNA      linear      EST 07-NOV-2001
LOCUS      60362203F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:5447542 5',
DEFINITION      mRNA sequence.
ACCESSION      BM044552
VERSION      BM044552.1 GI:16773819
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE      NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LNCM1931 row: j column: 23
High quality sequence stop: 639.
Location/Qualifiers
1..643
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5447542"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/notes="Organ: prostate; Vector: pOT87, Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Scratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

BASE COUNT      153 a      185 c      127 g      178 t

Query Match      10.4%; Score 642; DB 12; Length 643;
Best Local Similarity 100.0%; Pred. No. 3.3e-129;
Matches 642; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      4529 GATCTATCTGTTGCTAGGCTGAGTGCAGTGGCGCATCTCGGCTCACTGCAACCTC 4588
Db      1 GATCTATCTGTTGCTAGGCTGAGTGCAGTGGCGCATCTCGGCTCACTGCAACCTC 60
Oy      4589 TGCCTCTGGGCTCAAGCAATTCCTCCACCTCGCTCCCAATAGCTGGAGTCACTGGC 4648
Db      61 TGCCTCTGGGCTCAAGCAATTCCTCCACCTCGCTCCCAATAGCTGGAGTCACTGGC 120
Oy      4649 ACAACCACTGCGCACTAATTTGTATTTTGTAGAGACAGGCTTACCATGTTG 4708

```

```

Db      121 ACAACCACTGCGCACTAATTTGTATTTTGTAGAGACAGGCTTACCATGTTG 180
Oy      4709 CCCAGCTGTGTCTCAACCTCTCGGCTCAAGCAATCTCTGCTCGGCTCCAAAGTG 4768
Db      181 CCCAGCTGTGTCTCAACCTCTCGGCTCAAGCAATCTCTGCTCGGCTCCAAAGTG 240
Oy      4769 CTGGATTTACAGATGTGAGCCACCGCATCCAGCCCCACACCTCATTTTATACCAATTAC 4828
Db      241 CTGGATTTACAGATGTGAGCCACCGCATCCAGCCCCACACCTCATTTTATACCAATTAC 300
Oy      4829 TGCCAGTAACTGTGAGCTTTTGTCTTCTCAACCCCTGCTGATCTGGAAGAGAGGAT 4888
Db      301 TGCCAGTAACTGTGAGCTTTTGTCTTCTCAACCCCTGCTGATCTGGAAGAGAGGAT 360
Oy      4889 TATGTTATAGCTTGTACAGACAGTCCCAAGTCAATTTTCTGCGCAAAAACCTCTTC 4948
Db      361 TATGTTATAGCTTGTACAGACAGTCCCAAGTCAATTTTCTGCGCAAAAACCTCTTC 420
Oy      4949 AAAAAATTAATGTAATCTTCAATTTGATTTCAATGAATTCACCTTGGAAATGACCGCTCAAC 5008
Db      421 AAAAAATTAATGTAATCTTCAATTTGATTTCAATGAATTCACCTTGGAAATGACCGCTCAAC 480
Oy      5009 TTGTTCACTGAGCAATTAATGAAGAAATTTATAGTCTCTTAATAGCGCTGATGCAAG 5068
Db      481 TTGTTCACTGAGCAATTAATGAAGAAATTTATAGTCTCTTAATAGCGCTGATGCAAG 540
Oy      5069 ACCCTTGAACAATTCCAGAGATAGATATTTAAGTCATGCCCTTGCGCTATG 5128
Db      541 ACCCTTGAACAATTCCAGAGATAGATATTTAAGTCATGCCCTTGCGCTATG 600
Oy      5129 GCACCTTTCCTTCTGAAAAGTTGTTGTTCTGCTGCCAGTACC 5170
Db      601 GCACCTTTCCTTCTGAAAAGTTGTTGTTCTGCTGCCAGTACC 642

RESULT 11
CA307237/c      738 bp      mRNA      linear      EST 01-NOV-2002
LOCUS      UI-H-F11-Dhu-p-08-0-UI_g1 NCI CGAP F11 Homo sapiens cDNA clone
DEFINITION      UI-H-F11-Dhu-p-08-0-UI 3', mRNA sequence.
ACCESSION      CA307237
VERSION      CA307237.1 GI:24470291
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT      Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Dr. Gary W. Hunninghake, U of I
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=yes.
Location/Qualifiers
1..738
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-F11-Dhu-p-08-0-UI"
/tissue_type="Aveolar Macrophage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_F11"
/notes="Organ: Lung; Vector: pYT73-Pac (Pharmacia) with a

```

modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP FTL is a normalized cDNA library constructed from a pool of 81 RNA samples from Alveolar Macrophages challenged with different treatments. The library was normalized according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pTZ19-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GGGCATGGCG. The tissue was provided by Dr. Gary W. Hummshake of the University of Iowa.

TAG LIB=UI-H-FTL
TAG_TISSUE=Human Lung Alveolar Macrophage
TAG_SEQ=GGCATGGCG

BASE COUNT 145 a 195 c 229 g 168 t 1 others

ORIGIN

Query Match 10.3%; Score 635; DB 14; Length 738;
Best Local Similarity 100.0%; Pred. No. 9.7e-128; Indels 0; Gaps 0;
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1726 TACCTGTCGAGCCAGTCCATACCCAGCCAGCCTTCTGCTTCCCGGGA 1785
648 TACCTGTCGAGCCAGTCCATACCCAGCCAGCCTTCTGCTTCCCGGGA 589
1786 GGCTCTATCAAGCAGATCCAGGTGAAGCAAAATCTCGGTGACCTTGGACCTTTGGC 1845
588 GGCTCTATCAAGCAGATCCAGGTGAAGCAAAATCTCGGTGACCTTGGACCTTTGGC 529
1846 CCAGTTCACCAAGAGGCTCCAGGAGGCTGACGCTGCTTCTTACCTTATTTTC 1905
528 CCAGTTCACCAAGAGGCTCCAGGAGGCTGACGCTGCTTCTTACCTTATTTTC 469
1906 AAGAGAGAGGGCTTTTCAAGGTGACCCCTGACACAAAAGCAAGGTCTACCTGAGAGCC 1965
468 AAGAGAGAGGGCTTTTCAAGGTGACCCCTGACACAAAAGCAAGGTCTACCTGAGAGCC 409
1966 CCCAAGTGGAGCGGGGCTCCAGGAGGCTGACGCTGCTTCTTACCTTATTTTC 2025
408 CCCAAGTGGAGCGGGGCTCCAGGAGGCTGACGCTGCTTCTTACCTTATTTTC 349
2026 AAGAGAGAGGGCTTTTCAAGGTGACCCCTGACACAAAAGCAAGGTCTACCTGAGAGCC 2085
348 AAGAGAGAGGGCTTTTCAAGGTGACCCCTGACACAAAAGCAAGGTCTACCTGAGAGCC 289
2086 GGGCGGCAATTCATGATCAATCCAGAGCAGCGGAGCCGGGCTGAGAGATCTTCAAGCTG 2145
288 GGGCGGCAATTCATGATCAATCCAGAGCAGCGGAGCCGGGCTGAGAGATCTTCAAGCTG 229
2146 GAGAGAGATGCTCCCAAGCCAGCTTCCAGCATCAGAGCTTCTGGGTCAACATCTCT 2205
228 GAGAGAGATGCTCCCAAGCCAGCTTCCAGCATCAGAGCTTCTGGGTCAACATCTCT 169
2206 AACTGCAAGCCCAAGCGGCAAGCAGCTGACCTTCTCTGGGTGACATTAACCCA 2265
168 AACTGCAAGCCCAAGCGGCAAGCAGCTGACCTTCTCTGGGTGACATTAACCCA 109
2266 AAGAGAGATGCTGATCCTCATCGAGCGGCTGAGAGGTGAGATCTTACCTGCTG 2325
108 AAGAGAGATGCTGATCCTCATCGAGCGGCTGAGAGGTGAGATCTTACCTGCTG 49
2326 TCTGCTTCGGGCTCATCTTCTGTGTGAAAA 2360
48 TCTGCTTCGGGCTCATCTTCTGTGTGAAAA 14

RESULT 12
BG419603 851 bp mRNA linear EST 14-MAR-2001
LOCUS BG419603

DEFINITION 602451816F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4590554 5', mRNA sequence.
ACCESSION BG419603
VERSION BG419603.1 GI:13326109
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 851)
NIH-MGC <http://mgi.mgi.nhl.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LHCN1325 row: 0 column: 03
High quality sequence stop: 699.
Location/Qualifiers

FEATURES

source

1..851
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4590554"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 14"
/note="Organ: kidney; Vector: pOTB1; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 181 a 236 c 240 g 194 t

Query Match 10.2%; Score 629; DB 10; Length 851;
Best Local Similarity 100.0%; Pred. No. 1.7e-126; Indels 0; Gaps 0;
Matches 629; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

194 GGGGAGCGCAGCTGCGCGGCTTGGAGCGCTGAGGCGCGCGCTCCCAAGCTGTTT 253
47 GGGGAGCGCAGCTGCGCGGCTTGGAGCGCTGAGGCGCGCGCTCCCAAGCTGTTT 106
254 CCCACCGAGGCGAGGCGCTCCGAGGTGATGCGCGGCTGAACTGCGGGGCTCTATCG 313
107 CCCACCGAGGCGAGGCGCTCCGAGGTGATGCGCGGCTGAACTGCGGGGCTCTATCG 166
314 CACTGCTAGGGGTTTCTGCTGCTGCTGCTGCGCGCGGCTCCGCGGCGGAGAGCTTTG 373
167 CACTGCTAGGGGTTTCTGCTGCTGCTGCTGCGCGCGGCTCCGCGGCGGAGAGCTTTG 226
374 AGATTGCTCTGCGCAGAGAAAGCAATTAAGTCTTAAGCTGGGAGCCCGAGCTC 433
227 AGATTGCTCTGCGCAGAGAAAGCAATTAAGTCTTAAGCTGGGAGCCCGAGCTC 286
434 TGGTGGCAAAACCTGTTACATGTCATTTCTAAAGACATTAACATGTTGTCATCA 493
287 TGGTGGCAAAACCTGTTACATGTCATTTCTAAAGACATTAACATGTTGTCATCA 346
494 AGTTGGAGAAAGATTAATCTTACCTTAAGTGGCAGAGCTTGAATCACTTTGCA 553
347 AGTTGGAGAAAGATTAATCTTACCTTAAGTGGCAGAGCTTGAATCACTTTGCA 406
554 TAGAGATCCAGAAATAATTAATGCTATGTCAGGCCCATGTCCTTTGGGAGGTTGAGC 613

|||||
DB 407 TTGAGATCCAGAAAAATATGACTGATGTCAGAGCCCATGCTCTTTTGCGAGGTTCCAC 466
OY 614 TTCAGCCTCGACATGTTGTTGCTTACCTCAACAGAACTTTTCATCTGGGATGTCAAG 673
DB 467 TTCAGCCTCGACATGTTGTTGCTTACCTCAACAGAACTTTTCATCTGGGATGTCAAG 526
OY 674 CTCATAGAGCATCGTTTAGAGCTCAGTTTTCATCTCCCTCGCTGAGGAGATCGTC 733
DB 527 CTCATAGAGCATCGTTTAGAGCTCAGTTTTCATCTCCCTCGCTGAGGAGATCGTC 586
OY 734 CGGTTAGAGCTGCCAGAGAGGATCATCTTCATCGAGCGGCGAATCGATGCCACCG 793
DB 587 CGGTTAGAGCTGCCAGAGAGGATCATCTTCATCGAGCGGCGAATCGATGCCACCG 646
OY 794 TGGTCAGATCGGACCTTCTGAGCAAT 822
DB 647 TGGTCAGATCGGACCTTCTGAGCAAT 675

RESULT 13
BI457122 683 bp mRNA linear EST 21-AUG-2001
LOCUS 603185367F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5258105 5',
DEFINITION mRNA sequence.
ACCESSION BI457122
VERSION BI457122.1 GI:15247778
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 683)
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1869 row: e column: 18
High quality sequence stop: 678.
Location/Qualifiers

FEATURES
source
1. 683
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5258105"
/tissue_type="epitheloid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH_MGC_42"
/note="Organ: pancreas; Vector: pOT77; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAAGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library. |"

BASE COUNT 170 a 205 c 171 g 137 t
ORIGIN

Query Match 10.2%; Score 628; DB 12; Length 683;
Best Local Similarity 99.9%; Pred. No. 3,3e-126;
Matches 678; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1103 GTGAGAGAGAGAGAGCGGGTTGAATATCATCCCGGGCTCCACCAACCCCGAGG 1162

|||||
DB 2 GTGAGAGAGAGAGAGCGGGTTGAATATCATCTCCGGGCTCCACCAACCCCGAG 61
OY 1163 TGTTCAGGTGGAGAGAGAGAGCGGCTGGGAAACATGGCGGAGAACTTCAACCTCTGCG 1222
DB 62 TGTTCAGGTGGAGAGAGAGAGCGGCTGGGAAACATGGCGGAGAACTTCAACCTCTGCG 121
OY 1223 AAGGCTGTGACCAAGATGCCCAAGTCCAGGGATCTCCGCTGAGTTCCAAATTTTGG 1282
DB 122 AAGGCTGTGACCAAGATGCCCAAGTCCAGGGATCTCCGCTGAGTTCCAAATTTTGG 181
OY 1283 TCCACATCCCAAAATGAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1342
DB 182 TCCACATCCCAAAATGAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 241
OY 1343 CCATGCTACTCAGTACAGAGCCAGCCGCTGCAACAGAGCCGAATTTGTCCTGAGCT 1402
DB 242 CCATGCTACTCAGTACAGAGCCAGCCGCTGCAACAGAGCCGAATTTGTCCTGAGCT 301
OY 1403 GTTTCGTGTGTAGAAATCTCGGACCTGAGTACCTGACCTCCAGATCTGAGCTGCA 1462
DB 302 GTTTCGTGTGTAGAAATCTCGGACCTGAGTACCTGACCTCCAGATCTGAGCTGCA 361
OY 1463 AACACAAATCT 1522
DB 362 AACACAAATCT 421
OY 1523 CCATAGCTGACAG 1582
DB 422 CCATAGCTGACAG 481
OY 1583 GTGACATCTCTCAGCTGCTGCTGAGAGTGCATGACTTCTCTGAAAGCTGTGTCCTCA 1642
DB 482 GTGACATCTCTCAGCTGCTGCTGAGAGTGCATGACTTCTCTGAAAGCTGTGTCCTCA 541
OY 1643 AGGACAGGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1702
DB 542 AGGACAGGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 601
OY 1703 AGCCTGCAACAG 1762
DB 602 AGCCTGCAACAG 661
OY 1763 ACTTCGGCTCTCTCTGCCC 1781
DB 662 ACTTCGGCTCTCTCTGCCC 680

RESULT 14
BQ018619/c 712 bp mRNA linear EST 27-MAR-2002
LOCUS BQ018619
DEFINITION UI-H-DH1-awu-1-06-0-UI-81 NCI CGAP_DH1 Homo sapiens cDNA clone
IMAGE:5823821 3', mRNA sequence.
ACCESSION BQ018619.1 GI:19753896
VERSION BQ018619.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 712)
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Seq primer: M13 FORWARD
POLYA=Yes.

BASE COUNT	174 a	163 c	192 g	181 t	2 others
ORIGIN					

Query Match	10.1%;	Score 621;	DB 12;	Length 712;
Best Local Similarity	100.0%;	Pred. No. 1e-124;		
Matches 621;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	5292	GACATTTCCCAACCTCCAGGCCCTTCCAGAGAGACTAGGGGCTCATTTCCACCCAA	5351
Db	633	GACATTTCCCAACCTCCAGGCCCTTCCAGAGAGACTAGGGGCTCATTTCCACCCAA	574
QY	5352	GGTGGGATTTGGCTTCTTAGGCTGGTACTTGTCACCATGCAACGACATTCATGTGGCT	5411
Db	573	GGTGGGATTTGGCTTCTTAGGCTGGTACTTGTCACCATGCAACGACATTCATGTGGCT	514
QY	5412	GCAAGACACCAACGCGGACATTTTCCTTCAACTGAGGGCTCAAACTCTTGACAAAGTTG	5471
Db	513	GCAAGACACCAACGCGGACATTTTCCTTCAACTGAGGGCTCAAACTCTTGACAAAGTTG	454
QY	5472	CTGGCTCCTTGAGACACAGTATTTCTCTGAGAGCTGTGCTCAGTGAAGAGGGCCACGCTGAG	5531
Db	453	CTGGCTCCTTGAGACACAGTATTTCTCTGAGAGCTGTGCTCAGTGAAGAGGGCCACGCTGAG	394
QY	5532	AACCTGGCTCTTTTCTTTTAAAGCCACAGGCCCACTTACATPAAACATTTACAGGGTCACT	5591
Db	393	AACCTGGCTCTTTTCTTTTAAAGCCACAGGCCCACTTACATPAAACATTTACAGGGTCACT	334
QY	5592	GGAACAGTGAAGTGCCATTTGTGTAAGCCTACTGACATGCGACGCCCATGTGCTCATCCAG	5651
Db	333	GGAACAGTGAAGTGCCATTTGTGTAAGCCTACTGACATGCGACGCCCATGTGCTCATCCAG	274
QY	5652	TGCTCTGCGCATGCTTACGAGAGAGGCCACGCCGATGACAGACTGTCTCTTAATGCTGTGCT	5711
Db	273	TGCTCTGCGCATGCTTACGAGAGAGGCCACGCCGATGACAGACTGTCTCTTAATGCTGTGCT	214
QY	5712	CATTGCAACAGAGGGGAAGGTCTCAAGGAAGAAGTCAATTGGGACAAAGCAAGCCCAACCG	5771
Db	213	CATTGCAACAGAGGGGAAGGTCTCAAGGAAGAAGTCAATTGGGACAAAGCAAGCCCAACCG	154
QY	5772	GACATGACCTTGTGTAAGGTTAGCAGAGTGTGTGTGTGATTTGCGACATGCTTCACTGGA	5831
Db	153	GACATGACCTTGTGTAAGGTTAGCAGAGTGTGTGTGTGATTTGCGACATGCTTCACTGGA	94
QY	5832	AATAATTATTCATTGACAGATCTTTTAGGTGGCATTTTATTCATTTCTGTGTCTTAA	5891

Db 93 AATAATTATTCATTGACGACTCTTTTAGTGGCATTTTATTCATTTCGTGCTTAA 34

Qy 5892 ATAAACAATGTACCAAAAA 5912

Db 33 ATAAACAATGTACCAAAAA 13

RESULT	15
LOCUS	B0889198
DEFINITION	B0889198 871 bp mRNA linear EST 16-AUG-2002
ACCESSION	AGNCOCOURT_8072366 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6084315
VERSION	B0889198
KEYWORDS	5' , mRNA sequence.
SOURCE	B0889198.1 GI:2281212
ORGANISM	EST.
	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	1 (bases 1 to 871)
AUTHORS	NIH-MGC http://mgs.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished
COMMENT	Contact: Robert Strausberg, Ph.D.

Plate: LLCM2315 row: 0 column: 04
High quality sequence stop: 707.

Query March	9.9%	Score 612;	DB 13;	Length 871;
Best Local Similarity	99.7%	Pred. No. 7.9e-123;		
Matches 782; Conservative	0;	Mismatches 1;	Indels 1;	Gaps 1;

QY	3983	GATPATCTCTTCATCTTAAAGATCCCTCATCATCTCGAAGAGCCTTTTGCCATGCAAGAC	4042
Db	1	GATPATCTCTTCATCTTAAAGATCCCTCATCATCTCGAAGAGCCTTTTGCCATGCAAGAC	60
QY	4043	AACATAGCCCAAGGTGGGATTAGAGCCAGAGACATCTTTGGGGTCTGTTATTTCTGCTTA	4102
Db	61	AACATAGCCCAAGGTGGGATTAGAGCCAGAGACATCTTTGGGGTCTGTTATTTCTGCTTA	120
QY	4103	CCACACCTTCTCTGCCATCTGCCACAGAGAGGCTTACAAAATGATCTGGCGCACAGGG	4162
Db	121	CCACACCTTCTCTGCCATCTGCCACAGAGAGGCTTACAAAATGATCTGGCGCACAGGG	180
QY	4163	ATGTTTGTGTTAGCTTGGGACTCTAAACATT-AAAAAAAACCCGATCAGAGAGTCTGG	4221

444 ACCCTGTTACATCGTCATTTCTAAAGACATATACATGTTGTCATCATAGTCGGAGA 503
DB 242 ACCCTGTTACATCGTCATTTCTAAAGACATATACATGTTGTCATCATAGTCGGAGA 301
QY 504 AAGAAATAGCTTACCTTAGCGCGAGAGCTCGAATCACTTGTGATGAGATCA 563
DB 302 AAGAAATAGCTTACCTTAGCGCGAGAGCTCGAATCACTTGTGATGAGATCA 361
QY 564 GAAAAATATGACTGTATGTCAGGCCATGTCCTTTGGGGAGGTTCAAGCTTCAGCCCTC 623
DB 362 GAAAAATATGACTGTATGTCAGGCCATGTCCTTTGGGGAGGTTCAAGCTTCAGCCCTC 421
QY 624 GACATGTTGTTGCTTACCTTCAACAGAACTTTCATCTGGAGTGTCAAGCTCATAGAG 683
DB 422 GACATGTTGTTGCTTACCTTCAACAGAACTTTCATCTGGAGTGTCAAGAGCTCATAGAG 481
QY 684 CATCGGTTAGAGCTGACGATTTTCATCCCTGCGCTGAGAGAGATGGGTCCGGGTGAGAG 743
DB 482 CATCGGTTAGAGCTGACGATTTTCATCCCTGCGCTGAGAGAGATGGGTCCGGGTGAGAG 541
QY 744 CTGCCAGACGAGAGTCACTCATCATCAGCGGCCGAAATGATGCCAGCTGGTCAAGAT 803
DB 542 CTGCCAGACGAGAGTCACTCATCATCAGCGGCCGAAATGATGCCAGCTGGTCAAGAT 601
QY 804 CGGAACCTTCTGAGCAATGAGCACTGTGTCCGGATCAAGATGCAAGAGAGTGAAT 863
DB 602 TGGAACTTCTGAGCAATGAGCACTGTGTCCGGATCAAGATGCAAGAGAGTGAAT 661
QY 864 GGGCTTACACCTCCAGTGTCCACCCGAGAAATGTCTCCGGTTCAGAGATTGCAAAACGG 923
DB 662 GGGCTTACACCTCCAGTGTCCACCCGAGAAATGTCTCCGGTTCAGAGATTGCAAAACGG 721
QY 924 CTTCATCTATAAACGTCTGTGATCATCGAGTCTGTGTTGAGGGTGAAGGCTCAGCAAC 983
DB 722 CTTCATCTATAAACGTCTGTGATCATCGAGTCTGTGTTGAGGGTGAAGGCTCAGCAAC 781
QY 984 CCTGATGTCTGCCAATCAACAGAGGCTTCCCTGAGATGAGTCACTGATCGTGCAGTT 1043
DB 782 CCTGATGTCTGCCAATCAACAGAGGCTTCCCTGAGATGAGTCACTGATCGTGCAGTT 841
QY 1044 TGTGTTCCGTCACATCGTGGGGCGAGCGTCTCTCAACTTCAACCTTCCCAACG 1103
DB 842 TGTGTTCCGTCACATCGTGGGGCGAGCGTCTCTCAACTTCAACCTTCCCAACG 901
QY 1104 TGAAGAGAGAGAGAGCGGGTGAATCTACATCCCGGCTCCACCAACCCGAGGT 1163
DB 902 TGAAGAGAGAGAGAGCGGGTGAATCTACATCCCGGCTCCACCAACCCGAGGT 961
QY 1164 GTTCAAGCTGAGAGCAAGAGGCTTGGAAATGATGCGGGAACTTCAACCTCTCTGCA 1223
DB 962 GTTCAAGCTGAGAGCAAGAGGCTTGGAAATGATGCGGGAACTTCAACCTCTCTGCA 1021
QY 1224 AAGCGTGTGACCAAGATGCCCAAGTCCAGGGATCCCTCCGCTGCACTTCAAGTTTGGT 1283
DB 1022 AAGCGTGTGACCAAGATGCCCAAGTCCAGGGATCCCTCCGCTGCACTTCAAGTTTGGT 1081
QY 1284 CCAACATCCACAAATGAAAGCAATTAATCTACGTGTGACTTGAATGATGAGCAGC 1343
DB 1082 CCAACATCCACAAATGAAAGCAATTAATCTACGTGTGACTTGAATGATGAGCAGC 1141
QY 1344 CATGTCACTACCATGAGCAGCGCCCGTCAACAGAGCCGCAAGTTTGTCCCTGGCTG 1403
DB 1142 CATGTCACTACCATGAGCAGCGCCCGTCAACAGAGCCGCAAGTTTGTCCCTGGCTG 1201
QY 1404 TTTTCGTGTCTTAAGATCTCGACCTGAGCTGAGCACTCAACCTGACATCTGCTCCAA 1463
DB 1202 TTTTCGTGTCTTAAGATCTCGACCTGAGCTGAGCACTCAACCTGACATCTGCTCCAA 1261
QY 1464 ACACAAATCTCTCTCTTGTGATGATGACACCTCTGTGGATGAATGTGGAAAAAAC 1523
DB 1262 ACACAAATCTCTCTCTTGTGATGATGACACCTCTGTGGATGAATGTGGAAAAAAC 1321
QY 1524 CATTAAGCTGACAGACACCGGTACTGTCAGAAAGAAATCTTACTCATCTCAGGTGCCAG 1583

DB 1322 CATTAAGCTGACAGACACCGGTACTGTCAGAAAGAAATCTTACTCATCTCAGGTGCCAG 1381
QY 1584 TGAACATCTCCACCTGCTGTGAGAGCTGATGACTTCTCTGGAAGCTGTGGTCCAA 1643
DB 1382 TGAACATCTCCACCTGCTGTGAGAGCTGATGACTTCTCTGGAAGCTGTGGTCCAA 1441
QY 1644 GGAAGGCTCAGCTGTGTGTGTGTGTCAGGCCAGAGAGCTGACAGCATACACAGAA 1703
DB 1442 GGAAGGCTCAGCTGTGTGTGTGTGTCAGGCCAGAGAGCTGACAGCATACACAGAA 1501
QY 1704 GGCCTGCAACACACAGCTTCACTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1763
DB 1502 GGCCTGCAACACACAGCTTCACTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1561
QY 1764 CTTGGGCTCTCTCTCCGCGAGGCTCTATCAAGAGATCCAGGTGAAGCAGAACTCTC 1823
DB 1562 CTTGGGCTCTCTCTCCGCGAGGCTCTATCAAGAGATCCAGGTGAAGCAGAACTCTC 1621
QY 1824 GGTGACCTTGTGACCTTTGCCCCAGCTTCCAAAGAGGCTCTCAGGCAAGGTTGAC 1883
DB 1622 GGTGACCTTGTGACCTTTGCCCCAGCTTCCAAAGAGGCTCTCAGGCAAGGTTGAC 1681
QY 1884 GGTGTCTTTTATCTTATTTTCAAGAGAGAGGCTTTTCAAGGTGACCCCTGACACAA 1943
DB 1682 GGTGTCTTTTATCTTATTTTCAAGAGAGAGGCTTTTCAAGGTGACCCCTGACACAA 1741
QY 1944 AAGCAAGCTCTACCTGAGGACCCCACTGGGACCGGGGCTGCTCATCTCTCTGT 2003
DB 1742 AAGCAAGCTCTACCTGAGGACCCCACTGGGACCGGGGCTGCTCATCTCTCTGT 1801
QY 2004 GTCTGGAACATCAAGCTGTGCCAGAGACCAAGGTGCTGTCTGATCTTTTAAAGAGCG 2063
DB 1802 GTCTGGAACATCAAGCTGTGCCAGAGACCAAGGTGCTGTCTGATCTTTTAAAGAGCG 1861
QY 2064 GAGCGGCTGTGTCTTCCAGACAGAGGCTGCTGATGATGATGATGATGATGATGATG 2123
DB 1862 GAGCGGCTGTGTCTTCCAGACAGAGGCTGCTGATGATGATGATGATGATGATGATG 1921
QY 2124 GGTGAGAGATCTTCAAGCTGTGAGAGAGATGTGTCTCCCAAGCCAGGCTTCCACATCA 2183
DB 1922 GGTGAGAGATCTTCAAGCTGTGAGAGAGATGTGTCTCCCAAGCCAGGCTTCCACATCA 1981
QY 2184 CAGCTTCTGGGTCAACATCTTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2243
DB 1982 CAGCTTCTGGGTCAACATCTTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2041
QY 2244 CTTTCTGGGTGACATTTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2303
DB 2042 CTTTCTGGGTGACATTTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2101
QY 2304 GGGAGGTGAGCTTACTGCTGTGTGCTGCGGCTGATCATTTGCTGTGAAAAA 2360
DB 2102 GGGAGGTGAGCTTACTGCTGTGTGCTGCGGCTGATCATTTGCTGTGAAAAA 2158

RESULT 2
US-09-328-111-370/C
; Sequence 370 Application US/09328111
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroli III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dertl, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Nonahan, John E.
; APPLICANT: Schlegel, Robert

TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
FILE OF INVENTION: PRODUCTS
FILE REFERENCE: CCD-257 (US)
CURRENT APPLICATION NUMBER: US/09/328,111
CURRENT FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: US 60/088,801
EARLIER FILING DATE: 1998-06-10
NUMBER OF SEQ ID NOS: 850
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 370
LENGTH: 636
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)...(636)
OTHER INFORMATION: n = A,T,C or G
US-09-328-111-370

Query Match 3.8%; Score 236; DB 3; Length 636;
Best Local Similarity 99.7%; Pred. No. 3.6e-93;
Matches 286; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5617 AAGCTTACTGTCATGCCAGCCCTGCTCATCCAGTGTGCTCCATGCTACGAGAAAG 5676
DB 287 AAGCTTACTGTCATGCCAGCCCTGCTCATCCAGTGTGCTCCATGCTACGAGAAAG 228
OY 5677 CCAGGCATGCAAGCATGCTGCTCTAATGCTGTCATGCAAGAGGAAAGTCTCA 5736
DB 227 CCAGGCATGCAAGCATGCTGCTCTAATGCTGTCATGCAAGAGGAAAGTCTCA 168
OY 5737 AGGAAGATCACTGAGCAAGCAAGCCCAAGCCGACATGCTGTAAGGTTAGCA 5796
DB 167 AGGAAGATCACTGAGCAAGCAAGCCCAAGCCGACATGCTGTAAGGTTAGCA 108
OY 5797 GACTGTGTGTGATGATGTCAGTCTTCACTGAAATATTATTATTCATGAGATCTT 5856
DB 107 GACTGTGTGTGATGATGTCAGTCTTCACTGAAATATTATTATTCATGAGATCTT 48
OY 5857 TTTAGTGGCAATTTATTCATTTCCGTGCTTTAATAAACAATGT 5903
DB 47 TTTAGTGGCAATTTATTCATTTCCGTGCTTTAATAAACAATGT 1

RESULT 3

US-09-313-294A-7155
Sequence 7155, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Laligudi, Raghunath V.
APPLICANT: Ito, Laura Y.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 7155
LENGTH: 174
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. 6476212 700381216H1
US-09-313-294A-7155

Query Match 2.8%; Score 174; DB 4; Length 174;
Best Local Similarity 100.0%; Pred. No. 4e-66;
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1515 GAAAAAACATTAAGCTGACAGACACCGGTACTGCAAGAAAGAAATCTACTCACTCA 1574
|||||

DB 1 GAAAAAACATTAAGCTGACAGACACCGGTACTGCAAGAAAGAAATCTACTCACTCA 60
OY 1575 GATGCCAGTGCATCTCCACTGCTGATGAGCTGATGATCTTCTCTGGAAGTGT 1634
DB 61 GATGCCAGTGCATCTCCACTGCTGATGAGCTGATGATCTTCTCTGGAAGTGT 120
OY 1635 GGTGCCAAGGACAGCTCAGCTGTGCTGTGTCAGCCCGAAGCTGCAGCA 1688
DB 121 GGTGCCAAGGACAGCTCAGCTGTGCTGTGTCAGCCCGAAGCTGCAGCA 174

RESULT 4

US-09-292-542A-1/c
Sequence 1, Application US/09292542A
Patent No. 6531279
GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marta
APPLICANT: Chumakov, Ilya
APPLICANT: Bouguetelret, Lydie
TITLE OF INVENTION: Genomic Sequence Of The 5-Lipoxygenase-Activating Protein (FLAP),
Patent No. 6531279
FILE OF INVENTION: Polymorphic Markers Thereof And Methods For Detection Of Asthma.
FILE REFERENCE: GENSET.026A
CURRENT FILING DATE: 1999-04-15
CURRENT APPLICATION NUMBER: US/09/292,542A
PRIOR APPLICATION NUMBER: US 60/081893
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: US 60/091314
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/123406
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patent.pm
SEQ ID NO 1
LENGTH: 43069
TYPE: DNA
ORGANISM: homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1..7708
OTHER INFORMATION: potential 5' regulatory region
NAME/KEY: misc.feature
LOCATION: 36604..43069
OTHER INFORMATION: potential 3' regulatory region
NAME/KEY: exon
LOCATION: 7709..7852
OTHER INFORMATION: exon1
NAME/KEY: exon
LOCATION: 16236..16335
OTHER INFORMATION: exon2
NAME/KEY: exon
LOCATION: 24227..24297
OTHER INFORMATION: exon3
NAME/KEY: exon
LOCATION: 28133..28214
OTHER INFORMATION: exon4
NAME/KEY: exon
LOCATION: 36128..36605
OTHER INFORMATION: exon5
NAME/KEY: misc.feature
LOCATION: 7783..7785
OTHER INFORMATION: ATG
NAME/KEY: misc.feature
LOCATION: 36286..36290
OTHER INFORMATION: stop : TAA
NAME/KEY: POLYA signal
LOCATION: 36581..36586
OTHER INFORMATION: AATAAA
NAME/KEY: misc.feature
LOCATION: 7008..8116
OTHER INFORMATION: homology with sequence in ref genbank : M60470
NAME/KEY: misc.feature
LOCATION: 15995..16549

OTHER INFORMATION: homology with sequence in ref genbank : M63259
NAME/KEY: misc_feature
LOCATION: 24059..24597
OTHER INFORMATION: homology with sequence in ref genbank : M63260
NAME/KEY: misc_feature
LOCATION: 27875..28412
OTHER INFORMATION: homology with sequence in ref genbank : M63261
NAME/KEY: misc_feature
LOCATION: 35977..36926
OTHER INFORMATION: homology with sequence in ref genbank : M63262
NAME/KEY: misc_feature
LOCATION: 7613_
OTHER INFORMATION: diverging nucleotide deletion of a A in ref : M60470
NAME/KEY: misc_feature
LOCATION: 16347
OTHER INFORMATION: diverging nucleotide G in ref : M63259
NAME/KEY: misc_feature
LOCATION: 16348
OTHER INFORMATION: diverging nucleotide A in ref : M63259
NAME/KEY: misc_feature
LOCATION: 24060
OTHER INFORMATION: diverging nucleotide deletion of a G in ref : M63260
NAME/KEY: misc_feature
LOCATION: 24067
OTHER INFORMATION: diverging nucleotide deletion of a G in ref : M63260
NAME/KEY: misc_feature
LOCATION: 27903
OTHER INFORMATION: diverging nucleotide deletion of a C in ref : M63261
NAME/KEY: misc_feature
LOCATION: 28327
OTHER INFORMATION: diverging nucleotide deletion of a G in ref : M63261
NAME/KEY: misc_feature
LOCATION: 3851..4189
OTHER INFORMATION: 10-517
NAME/KEY: misc_feature
LOCATION: 4120..4390
OTHER INFORMATION: 10-518
NAME/KEY: misc_feature
LOCATION: 4373..4792
OTHER INFORMATION: 10-253
NAME/KEY: misc_feature
LOCATION: 4814..5043
OTHER INFORMATION: 10-499
NAME/KEY: misc_feature
LOCATION: 4956..5422
OTHER INFORMATION: 10-500
NAME/KEY: misc_feature
LOCATION: 5524..5996
OTHER INFORMATION: 10-522
NAME/KEY: misc_feature
LOCATION: 6218..6672
OTHER INFORMATION: 10-503
NAME/KEY: misc_feature
LOCATION: 6522..6790
OTHER INFORMATION: 10-504
NAME/KEY: misc_feature
LOCATION: 7120..7574
OTHER INFORMATION: 10-204
NAME/KEY: misc_feature
LOCATION: 7513..7933
OTHER INFORMATION: 10-32
NAME/KEY: misc_feature
LOCATION: 16114..16533
OTHER INFORMATION: 10-33
NAME/KEY: misc_feature
LOCATION: 24072..24425
OTHER INFORMATION: 10-34
NAME/KEY: misc_feature
LOCATION: 27978..28401
OTHER INFORMATION: 10-35
NAME/KEY: misc_feature
LOCATION: 36020..36465
OTHER INFORMATION: 10-36

NAME/KEY: misc_feature
LOCATION: 36318..36669
OTHER INFORMATION: 10-498
NAME/KEY: misc_feature
LOCATION: 38441..38840
OTHER INFORMATION: 12-629
NAME/KEY: misc_feature
LOCATION: 42235..42749
OTHER INFORMATION: 12-628 complement
NAME/KEY: allele
LOCATION: 3950
OTHER INFORMATION: 10-517-100 : polymorphic base S
NAME/KEY: allele
LOCATION: 4243
OTHER INFORMATION: 10-518-125 : polymorphic base K
NAME/KEY: allele
LOCATION: 4312
OTHER INFORMATION: 10-518-194 : polymorphic base R
NAME/KEY: allele
LOCATION: 4490
OTHER INFORMATION: 10-253-118 : polymorphic base R
NAME/KEY: allele
LOCATION: 4670
OTHER INFORMATION: 10-253-298 : polymorphic base S
NAME/KEY: allele
LOCATION: 4687
OTHER INFORMATION: 10-253-315 : polymorphic base Y
NAME/KEY: allele
LOCATION: 4968
OTHER INFORMATION: 10-499-155 : polymorphic base R
NAME/KEY: allele
LOCATION: 5140
OTHER INFORMATION: 10-500-185 : polymorphic base Y
NAME/KEY: allele
LOCATION: 5213
OTHER INFORMATION: 10-500-258 : polymorphic base K
NAME/KEY: allele
LOCATION: 5364
OTHER INFORMATION: 10-500-410 : polymorphic base R
NAME/KEY: allele
LOCATION: 5594
OTHER INFORMATION: 10-522-71 : polymorphic base R
NAME/KEY: allele
LOCATION: 6370
OTHER INFORMATION: 10-503-159 : polymorphic base K
NAME/KEY: allele
LOCATION: 6693
OTHER INFORMATION: 10-504-172 : polymorphic base W
NAME/KEY: allele
LOCATION: 6763
OTHER INFORMATION: 10-504-243 : polymorphic base M
NAME/KEY: allele
LOCATION: 7445
OTHER INFORMATION: 10-204-326 : polymorphic base R
NAME/KEY: allele
LOCATION: 7870
OTHER INFORMATION: 10-32-357 : polymorphic base M
NAME/KEY: allele
LOCATION: 16288
OTHER INFORMATION: 10-33-175 : polymorphic base Y
NAME/KEY: allele
LOCATION: 16347
OTHER INFORMATION: 10-33-234 : polymorphic base M
NAME/KEY: allele
LOCATION: 16383
OTHER INFORMATION: 10-33-270 : polymorphic base R
NAME/KEY: allele

Query Match . 0.9% ; Score 53 ; DB 4 ; Length 43069 ;
Best Local Similarity 100.0% ; Pred. No. 1.8e-13 ;
Matches 53 ; Conservative 0 ; Mismatches 0 ; Indels 0 ; Gaps 0 ;

Qy 4547 AGCTGAGTGCAGTGGCGCAATCTGGCTACTGCAACTCTGCTCTGCG 4599

Db 28976 AGCGTGAAGTGCAGCGCCCAATCTCGGCTCACTGCAACCTCTGCTCTGGG 28924

US-09-146-053-6/c
Sequence 6, Application US/09146053A
Patent No. 6399349

GENERAL INFORMATION:

APPLICANT: Ryan, James W.

APPLICANT: Sprinkle, Terry Joe Curtis

APPLICANT: Venema, Richard C.

FILE REFERENCE: MCG103

CURRENT FILING DATE: 1998-09-02

EARLIER APPLICATION NUMBER: 60/057,854

NUMBER OF SEQ ID NOS: 7

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 6

LENGTH: 45546

TYPE: DNA

ORGANISM: Homo sapiens

US-09-146-053-6

Query Match 0.8%; Score 51; DB 4; Length 45546;

Best Local Similarity 100.0%; Pred. No. 1.3e-12;

Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

4671 TTTTGTATTTTGTAGAGACAGAGGTTTCACCATGTTGCCAGGCTGTCT 4721

DB 39160 TTTTGTATTTTGTAGAGACAGAGGTTTCACCATGTTGCCAGGCTGTCT 39110

RESULT 6

US-09-851-896-3

Sequence 3, Application US/09851896

Patent No. 6410325

GENERAL INFORMATION:

APPLICANT: C. Frank Bennett

APPLICANT: Susan M. Preler

FILE REFERENCE: RTS-0220

CURRENT FILING DATE: 2001-05-08

NUMBER OF SEQ ID NOS: 89

SEQ ID NO 3

LENGTH: 70000

TYPE: DNA

ORGANISM: Homo sapiens

US-09-851-896-3

Query Match 0.8%; Score 49; DB 4; Length 70000;

Best Local Similarity 100.0%; Pred. No. 9.5e-12;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

4547 AGCGTGAAGTGCAGCGCCCAATCTCGGCTCACTGCAACCTCTGCTCTCC 4595

DB 11896 AGCGTGAAGTGCAGCGCCCAATCTCGGCTCACTGCAACCTCTGCTCTCC 11944

RESULT 7

US-08-994-076-1/c

Sequence 1, Application US/08994076

Patent No. 6500937

GENERAL INFORMATION:

APPLICANT: Ervin, Jr., Paul R.

FILE REFERENCE: NUCLEOTIDE AND PROTEIN SEQUENCE

TITLE OF INVENTION: OF MAMMASTATIN AND METHODS OF USE

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESS: Merchant, Gould, Smith, Edell, Welter & Schmidt

STREET: 3100 No. 6500937 West Center, 90 South 7th Street

CITY: Minneapolis

STATE: MN

COUNTRY: USA

ZIP: 55402

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/994,076

FILING DATE: 19-DEC-1997

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/027,315

FILING DATE: 03-OCT-1996

APPLICATION NUMBER: PCT/US97/18026

FILING DATE: 03-OCT-1997

APPLICATION NUMBER: 08/943,828

FILING DATE: 03-OCT-1997

ATTORNEY/AGENT INFORMATION:

NAME: Ketelberger, Denise M

REGISTRATION NUMBER: 33,924

REFERENCE/DOCKET NUMBER: 4273.1US11

TELECOMMUNICATION INFORMATION:

TELEPHONE: 612-332-5100

TELEFAX: 612-332-9081

TELEX:

SEQUENCE CHARACTERISTICS:

LENGTH: 2326 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: Genomic DNA

FEATURE:

NAME/KEY: Coding Sequence

LOCATION: 348...1961

OTHER INFORMATION:

US-08-994-076-1

Query Match 0.7%; Score 45; DB 4; Length 2326;

Best Local Similarity 100.0%; Pred. No. 6e-10;

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

4678 TTTTGTAGAGACAGAGGTTTCACCATGTTGCCAGGCTGTCTC 4722

DB 2144 TTTTGTAGAGACAGAGGTTTCACCATGTTGCCAGGCTGTCTC 2100

RESULT 8

US-09-285-379-1/c

Sequence 1, Application US/09285379

Patent No. 6451765

GENERAL INFORMATION:

APPLICANT: Ervin Jr., Paul R.

FILE REFERENCE: METHODS OF USE

TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCE OF MAMMASTATIN AND

TYPE: DNA
ORGANISM: Homo sapiens
US-09-285-379-1

Query Match
Best Local Similarity 100.0%; Score 45; DB 4; Length 2418;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4678 TTTTGTAGAGACAGAGGTTTCACCATGTTGCCAGGCTGTCTC 4722
DB 2233 TTTTGTAGAGACAGAGGTTTCACCATGTTGCCAGGCTGTCTC 2189

RESULT 9
US-09-369-912-1/c
Sequence 1, Application US/09369912
Patent No. 6492504
GENERAL INFORMATION:
APPLICANT: Evin, Jr., Paul R.
TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCE OF MAMMASTATIN AND
FILE REFERENCE: 4273.1USU1
CURRENT APPLICATION NUMBER: US/09/369,912
EARLIER FILING DATE: 1999-08-06
EARLIER APPLICATION NUMBER: 60/027,315
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin Ver. 2.0
SEO ID NO 1
LENGTH: 2418
TYPE: DNA
ORGANISM: Homo sapiens
US-09-369-912-1

Query Match
Best Local Similarity 100.0%; Score 45; DB 4; Length 2418;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4678 TTTTGTAGAGACAGAGGTTTCACCATGTTGCCAGGCTGTCTC 4722
DB 2233 TTTTGTAGAGACAGAGGTTTCACCATGTTGCCAGGCTGTCTC 2189

RESULT 10
US-08-975-080-35/c
Sequence 35, Application US/08975080
Patent No. 6245523
GENERAL INFORMATION:
APPLICANT: Allieri, Dario C.
TITLE OF INVENTION: SURVIVIN, A PROTEIN THAT INHIBITS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
STREET: 1800 M Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5869
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/975,080
FILING DATE: 20-NOV-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/031,435
FILING DATE: 20-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Reid G.
REGISTRATION NUMBER: 30,988

REFERENCE/DOCKET NUMBER: 044574-5022-01-WO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-467-7000
TELEFAX: 202-467-7176
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 14796 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-975-080-35

Query Match
Best Local Similarity 100.0%; Score 45; DB 3; Length 14796;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4747 CCTGCTGGGCTCCCAAGTGTGGATTACAGATGAGCCAC 4791
DB 1410 CCTGCTGGGCTCCCAAGTGTGGATTACAGATGAGCCAC 1366

RESULT 11
US-09-630-706-10/c
Sequence 10, Application US/09630706
Patent No. 6277640
GENERAL INFORMATION:
APPLICANT: C. Frank Bennett
TITLE OF INVENTION: ANTISENSE MODULATION OF HER-3 EXPRESSION
FILE REFERENCE: RTS-0053
CURRENT APPLICATION NUMBER: US/09/630,706
CURRENT FILING DATE: 2000-08-01
NUMBER OF SEQ ID NOS: 94
SEO ID NO 10
LENGTH: 14796
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (2811)...(2921)
NAME/KEY: CDS
LOCATION: (3174)...(3283)
NAME/KEY: CDS
LOCATION: (5158)...(5275)
NAME/KEY: CDS
LOCATION: (11955)...(12044)
US-09-630-706-10

Query Match
Best Local Similarity 100.0%; Score 45; DB 3; Length 14796;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4747 CCTGCTGGGCTCCCAAGTGTGGATTACAGATGAGCCAC 4791
DB 1410 CCTGCTGGGCTCCCAAGTGTGGATTACAGATGAGCCAC 1366

RESULT 12
US-09-496-694B-3/c
Sequence 3, Application US/09496694B
Patent No. 6335194
GENERAL INFORMATION:
APPLICANT: C. Frank Bennett
APPLICANT: Elizabeth J. Ackermann
APPLICANT: Eric E. Swartz
TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION
FILE REFERENCE: ISPH-0439
CURRENT APPLICATION NUMBER: US/09/496,694B
CURRENT FILING DATE: 2000-02-02
PRIOR APPLICATION NUMBER: 09/286,407
PRIOR FILING DATE: 1999-04-05

PRIOR APPLICATION NUMBER: 09/163,162
PRIOR FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 249
SEQ ID NO 3
LENGTH: 14796
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (2811)...(2921)
NAME/KEY: CDS
LOCATION: (3174)...(3283)
NAME/KEY: CDS
LOCATION: (5158)...(5275)
NAME/KEY: CDS
LOCATION: (11955)...(12044)
US-09-496-694B-3

Query Match 0.7%; Score 45; DB 4; Length 14796;
Best Local Similarity 100.0%; Pred. No. 5.6e-10;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4747 CCTGCTCGGCTCCCAAGTGTGGATTACAGATGTAGCCAC 4791
DB 1410 CCTGCTCGGCTCCCAAGTGTGGATTACAGATGTAGCCAC 1366

RESULT 13
US-09-735-934A-3/C
Sequence 3, Application US/09735934A
Patent No. 6372468
GENERAL INFORMATION:
APPLICANT: Li, JiaYin et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
FILE REFERENCE: CL000851
CURRENT APPLICATION NUMBER: US/09/735,934A
CURRENT FILING DATE: 2000-12-14
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 43950
TYPE: DNA
ORGANISM: Homo sapiens
US-09-735-934A-3

Query Match 0.7%; Score 44; DB 4; Length 43950;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4678 TTTTGTAGAGACAGGGTTTCAACCATGTTGCCAGGCTGTCT 4721
DB 14702 TTTTGTAGAGACAGGGTTTCAACCATGTTGCCAGGCTGTCT 14659

RESULT 14
US-10-060-332-3/C
Sequence 3, Application US/10060332
Patent No. 6528294
GENERAL INFORMATION:
APPLICANT: Li, JiaYin et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
FILE REFERENCE: CL000851DIV
CURRENT APPLICATION NUMBER: US/10/060,332
CURRENT FILING DATE: 2002-02-01
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 43950
TYPE: DNA

ORGANISM: Homo sapiens
US-10-060-332-3

Query Match 0.7%; Score 44; DB 4; Length 43950;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4678 TTTTGTAGAGACAGGGTTTCAACCATGTTGCCAGGCTGTCT 4721
DB 14702 TTTTGTAGAGACAGGGTTTCAACCATGTTGCCAGGCTGTCT 14659

RESULT 15
US-09-128-155-17/C
Sequence 17, Application US/09128155
Patent No. 6117654
GENERAL INFORMATION:
APPLICANT: Pan, Yang
TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
FILE REFERENCE: 09404/052001
CURRENT APPLICATION NUMBER: US/09/128,155
CURRENT FILING DATE: 1998-08-03
EARLIER APPLICATION NUMBER: US 60/091,650
EARLIER FILING DATE: 1998-07-02
EARLIER APPLICATION NUMBER: US 60/054,646
EARLIER FILING DATE: 1997-08-04
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 17
LENGTH: 176373
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(176373)
OTHER INFORMATION: n = A,T,C or G
US-09-128-155-17

Query Match 0.7%; Score 44; DB 3; Length 176373;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4750 GCCTGGGCTCCCAAGTGTGGATTACAGATGTAGCCACCG 4793
DB 168946 GCCTGGGCTCCCAAGTGTGGATTACAGATGTAGCCACCG 168903

Search completed: February 20, 2004, 13:55:05
Job time : 340 secs

This Page Blank (uspto)


```

; LOCATION: (2794) ..(6163)
; NAME/KEY: CDS
; LOCATION: (283) ..(2793)
US-09-899-569A-3

```

Query Match	100.0%;	Score 6163;	DB 10;	Length 6163;
-------------	---------	-------------	--------	--------------

Matches 6163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CCAA	CGCCG	CAATGGG	AGTGTGGG	AGCCG	CAAC	CCCGT	TGCGGG	AGCC	TG	CA	CC	60	
Db	1	CCAA	CGCCG	CAATGGG	AGTGTGGG	AGCCG	CAAC	CCCGT	TGCGGG	AGCC	TG	CA	CC	60	
Qy	61	TGGG	AGGGA	GAGG	CGGT	GTG	TGA	GCA	GAA	GAG	GAG	GAG	GGA	CG	120
Db	61	TGGG	AGGGA	GAGG	CGGT	GTG	TGA	GCA	GAA	GAG	GAG	GAG	GGA	CG	120
Qy	121	CGGG	TCAG	CTG	CGCAT	CTG	CTG	CG	CA	AGG	AGG	AGG	AGG	CG	180
Db	121	CGGG	TCAG	CTG	CGCAT	CTG	CTG	CG	CA	AGG	AGG	AGG	AGG	CG	180
Qy	181	GTG	AGT	AG	CGAGG	CGG	AG	CG	CTG	CGG	CTT	G	GG	CG	240
Db	181	GTG	AGT	AG	CGAGG	CGG	AG	CG	CTG	CGG	CTT	G	GG	CG	240
Qy	241	CCAC	CGT	CGTTT	CCCA	CCG	AG	CG	GGG	GT	CC	CG	AG	TC	300
Db	241	CCAC	CGT	CGTTT	CCCA	CCG	AG	CG	GGG	GT	CC	CG	AG	TC	300
Qy	301	GGG	CTCT	TA	CGCA	CTG	CTA	AGG	GGT	TG	CTG	TG	CG	CG	360
Db	301	GGG	CTCT	TA	CGCA	CTG	CTA	AGG	GGT	TG	CTG	TG	CG	CG	360
Qy	361	GCAG	AG	CTTT	TG	AG	ATT	G	CTT	CG	CAC	GAG	AA	GC	420
Db	361	GCAG	AG	CTTT	TG	AG	ATT	G	CTT	CG	CAC	GAG	AA	GC	420
Qy	421	GGG	AC	CCG	CA	CTG	CTG	CG	CAAAA	CC	CGT	TA	CG	CA	480
Db	421	GGG	AC	CCG	CA	CTG	CTG	CG	CAAAA	CC	CGT	TA	CG	CA	480
Qy	481	ATG	TG	TC	CA	CA	AGT	CTG	AG	AA	GA	TA	GT	CTT	540
Db	481	ATG	TG	TC	CA	CA	AGT	CTG	AG	AA	GA	TA	GT	CTT	540
Qy	541	AAT	CA	CTTT	GC	AT	AG	AG	AT	TC	CA	AAAA	AT	TG	600
Db	541	AAT	CA	CTTT	GC	AT	AG	AG	AT	TC	CA	AAAA	AT	TG	600
Qy	601	GGG	AG	GT	TC	AG	CC	CT	GA	CA	TC	GT	TG	CC	660
Db	601	GGG	AG	GT	TC	AG	CC	CT	GA	CA	TC	GT	TG	CC	660
Qy	661	TGG	AT	GT	CA	AA	G	CT	CA	T	TA	G	AG	CT	720
Db	661	TGG	AT	GT	CA	AA	G	CT	CA	T	TA	G	AG	CT	720
Qy	721	AGG	CA	AT	CG	GT	CG	GA	G	CT	GC	CA	AG	CA	780
Db	721	AGG	CA	AT	CG	GT	CG	GA	G	CT	GC	CA	AG	CA	780
Qy	781	ATC	AT	GC	CA	CC	GT	GC	AG	AT	CG	AA	CT	TG	840
Db	781	ATC	AT	GC	CA	CC	GT	GC	AG	AT	CG	AA	CT	TG	840
Qy	841	AAG	AT	GC	CA	AG	AG	AT	GA	AA	T	GC	CTT	CA	900
Db	841	AAG	AT	GC	CA	AG	AG	AT	GA	AA	T	GC	CTT	CA	900
Qy	901	TCG	CG	CTT	CA	GC	AT	TG	CA	AA	CG	GT	CA	T	960
Db	901	TCG	CG	CTT	CA	GC	AT	TG	CA	AA	CG	GT	CA	T	960

QY	961	TTTGAGGGTGAAGGCTCAGCAACCCTGATGTGGCCACTACCCGAAGAAGGCTTCCCTAG	1020
Db	961	TTTGAGGGTGAAGGCTCAGCAACCCTGATGTGGCCACTACCCGAAGAAGGCTTCCCTAG	1020
QY	1021	GATGAGCTCAATGACGTGGCAGTTTGTCTGTTCCGACACCTGCGGGCCAGGCTCTCTTC	1080
Db	1021	GATGAGCTCAATGACGTGGCAGTTTGTCTGTTCCGACACCTGCGGGCCAGGCTCTCTTC	1080
QY	1081	CTCAACTTCAACTCTTCACACTGTGTAGAGGAAGAGAGCGGGTTGAATACTACTACCTCCG	1140
Db	1081	CTCAACTTCAACTCTTCACACTGTGTAGAGGAAGAGAGCGGGTTGAATACTACTACCTCCG	1140
QY	1141	GGCTCCACCCCAACCCCGAGGGTTCACAGCTGAGAGGACAAGCAGCCTGGGAAACATGGCG	1200
Db	1141	GGCTCCACCCCAACCCCGAGGGTTCACAGCTGAGAGGACAAGCAGCCTGGGAAACATGGCG	1200
QY	1201	GGGAATTTCAACCTCTCTCTGCAAGGCTGTGACCAAGATGCCCAAGTCCAGAGATCTTC	1260
Db	1201	GGGAATTTCAACCTCTCTCTGCAAGGCTGTGACCAAGATGCCCAAGTCCAGAGATCTTC	1260
QY	1261	CGGCTGCAGTTCCAAAGTTTTGGTCCAACTGCACAAAAATGAAGCATPAAATCTACGTG	1320
Db	1261	CGGCTGCAGTTCCAAAGTTTTGGTCCAACTGCACAAAAATGAAGCATPAAATCTACGTG	1320
QY	1321	GTTGACTTGAATGATAGGAGGCGCATATGTACTACCATATGAGCCACCGGCCCTGCAAAAG	1380
Db	1321	GTTGACTTGAATGATAGGAGGCGCATATGTACTACCATATGAGCCACCGGCCCTGCAAAAG	1380
QY	1381	AGCCGCAAGTTTGTCTCCCTGGCTGTTTCGTGTCTAGAAATCTCGGACCTGCAATGCAAC	1440
Db	1381	AGCCGCAAGTTTGTCTCCCTGGCTGTTTCGTGTCTAGAAATCTCGGACCTGCAATGCAAC	1440
QY	1441	CTCAACCTGCACATCTGGCTCAAAACAACAAATCTCTCTTGTGATGATCTGCACAGT	1500
Db	1441	CTCAACCTGCACATCTGGCTCAAAACAACAAATCTCTCTTGTGATGATCTGCACAGT	1500
QY	1501	CTGTGATGAAATGTGGAAAAAACATTAAGCTGCACAGACCAACCGGTATCTGGCAAAAGGAA	1560
Db	1501	CTGTGATGAAATGTGGAAAAAACATTAAGCTGCACAGACCAACCGGTATCTGGCAAAAGGAA	1560
QY	1561	TTCCTACTCACTCCAGGTCGCAAGTGACATCTCTGCACTGCGCTGTGAGAGCTGATGACTTC	1620
Db	1561	TTCCTACTCACTCCAGGTCGCAAGTGACATCTCTGCACTGCGCTGTGAGAGCTGATGACTTC	1620
QY	1621	TCTGTGAAGCTGTGTGTGCCCAAGACAGGCTCAGCCTGTGTCTGTCACAGCCAGAG	1680
Db	1621	TCTGTGAAGCTGTGTGTGCCCAAGACAGGCTCAGCCTGTGTCTGTCACAGCCAGAG	1680
QY	1681	CTGACAGCAGATACACACAGAAAGCCCTGCAACACACAGTTCACTGCTGTGGCCAGT	1740
Db	1681	CTGACAGCAGATACACACAGAAAGCCCTGCAACACACAGTTCACTGCTGTGGCCAGT	1740
QY	1741	GCAATTAACCAAGCAAGGACTGTAACTTTGGGCTCTCTGCGCGGAGGCTTATACAGAG	1800
Db	1741	GCAATTAACCAAGCAAGGACTGTAACTTTGGGCTCTCTGCGCGGAGGCTTATACAGAG	1800
QY	1801	ATCCAGGTGAAGAGAAACATCTCGGTGACCTTCTGCACTTTGCCCCCAGCTTCAACAA	1860
Db	1801	ATCCAGGTGAAGAGAAACATCTCGGTGACCTTCTGCACTTTGCCCCCAGCTTCAACAA	1860
QY	1861	GAGGCTTCCAGGAGGGTGTGACGSGTGTCTTTATACCTTATTTCAAAAGAGAGCGTT	1920
Db	1861	GAGGCTTCCAGGAGGGTGTGACGSGTGTCTTTATACCTTATTTCAAAAGAGAGCGTT	1920
QY	1921	TTTACAGGTGACCCCTTACACAAAAAGCAAGGCTTACCTGAGAGACCCCACTGGAGCCGG	1980
Db	1921	TTTACAGGTGACCCCTTACACAAAAAGCAAGGCTTACCTGAGAGACCCCACTGGAGCCGG	1980
QY	1981	GGGCTGCACATCCCTTACCTCTGTGTGTCGAGAAACATCAGAGTGCCCAAGAGCCAGGTGGCC	2040
Db	1981	GGGCTGCACATCCCTTACCTCTGTGTGTCGAGAAACATCAGAGTGCCCAAGAGCCAGGTGGCC	2040
QY	2041	TGCTCACTTTCTTTTAAGAGCGGAGCGGCGTGTCTGCGACAGAGGCGCGCATTTGATG	2100

```
|||||
Db 2041 TGCGTCATCTTTTAAAGAGCGGAGCGCGTGTCTCCAGACAGCGCGCATTCATG 2100
Qy ATCATTCAGAGAGCGCGGCTGAGAGATCTTCAGCGTGGAGAGAGATGTCTC 2160
Db 2101 ATCATTCAGAGAGCGGAGCGCGTGTGAGAGATCTTCAGCGTGGAGAGATGTCTC 2160
Qy 2161 CCCAAGCCAGACTTCACACATCACAGCTTCTGGGTCAACATCTCTAATCTGAGCCCAAG 2220
Db 2161 CCCAAGCCAGACTTCACACATCACAGCTTCTGGGTCAACATCTCTAATCTGAGCCCAAG 2220
Qy 2221 AGCGGCAAGAGCTAGACCTGTCTCTTCTGGGTGACATTTACCCCAAGGACTGTGAGCTTG 2280
Db 2221 AGCGGCAAGAGAGCTAGACCTGTCTCTTCTGGGTGACATTTACCCCAAGGACTGTGAGCTTG 2280
Qy 2281 ACTGTCAATCCATATCGAGCGGTGGAGGTGAGTCTTACTGTCTGTCTCGGCTC 2340
Db 2281 ACTGTCAATCCATATCGAGCGGTGGAGGTGAGTCTTACTGTCTGTCTCGGCTC 2340
Qy 2341 ATCATTTGCTGTGTGAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2400
Db 2341 ATCATTTGCTGTGTGAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2400
Qy 2401 TACAAATGGCAACATCAATATCTGAGATGCGAGGAGCCAAAAAGTTTCAAGAAAGGCGA 2460
Db 2401 TACAAATGGCAACATCAATATCTGAGATGCGAGGAGCCAAAAAGTTTCAAGAAAGGCGA 2460
Qy 2461 AAGGACAAATGATCCCATATGTATGATGATCCAGAGACCAATGATATATGGGATCTG 2520
Db 2461 AAGGACAAATGATCCCATATGTATGATGATCCAGAGACCAATGATATATGGGATCTG 2520
Qy 2521 CTACAGATTCACAGCGCTCTCTCTCTGAGCGAGAGGTGACACCTACCGGCGTTCCAG 2580
Db 2521 CTACAGATTCACAGCGCTCTCTCTCTGAGCGAGAGGTGACACCTACCGGCGTTCCAG 2580
Qy 2581 GGCAACATGGGGGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2640
Db 2581 GGCAACATGGGGGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2640
Qy 2641 AAGTTGGGCACTGAGAGAGCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2700
Db 2641 AAGTTGGGCACTGAGAGAGCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2700
Qy 2701 ACCTTCTCCATCCCAACATGGGGATGTAAAGACAGAGACACAGACATTCCTTACTG 2760
Db 2701 ACCTTCTCCATCCCAACATGGGGATGTAAAGACAGAGACACAGACATTCCTTACTG 2760
Qy 2761 AACACTCAGAGAGCCCATGAGAGCCAGACAGATTAATCTGATTCATTCAGACGCTTGTGA 2820
Db 2761 AACACTCAGAGAGCCCATGAGAGCCAGACAGATTAATCTGATTCATTCAGACGCTTGTGA 2820
Qy 2821 GTTTCATTAAGAGAGGCACTGAGACACCGGTCTGTCTCTCTCTCTCTCTCTCTCTCTCT 2880
Db 2821 GTTTCATTAAGAGAGGCACTGAGACACCGGTCTGTCTCTCTCTCTCTCTCTCTCTCTCT 2880
Qy 2881 GAGGAATTAATCAGAAAGAAAGAGAGAGTTTCTCTGAGACACCGGCACTTCAATGCG 2940
Db 2881 GAGGAATTAATCAGAAAGAAAGAGAGAGTTTCTCTGAGACACCGGCACTTCAATGCG 2940
Qy 2941 TCAGTGACATCATTTAAAGGCAAGACATTTGAAATGATGATTCATTCGATATACAGT 3000
Db 2941 TCAGTGACATCATTTAAAGGCAAGACATTTGAAATGATGATTCATTCGATATACAGT 3000
Qy 3001 CATGACAGCTCATGTGCTCTCTCAATTAAGGCTGTGCGGTGAGCCAGCTGTATATAGAG 3060
Db 3001 CATGACAGCTCATGTGCTCTCTCAATTAAGGCTGTGCGGTGAGCCAGCTGTATATAGAG 3060
Qy 3061 AGAGAGGCTAGTCACTTAGATAGGTGAGAGAGCCCTGATTCAGAGTGTAAAC 3120
Db 3061 AGAGAGGCTAGTCACTTAGATAGGTGAGAGAGCCCTGATTCAGAGTGTAAAC 3120
Qy 3121 AGAGGCTTGCCTCTTCAAGAGCAACAGTTCCAAATTCAGAGAGCTTACCTGAGTCCCTA 3180
|||||

Db 3121 AGAGGCTTGCCTCTTCAAGAGCAACAGTTCCAAATTCAGAGAGCTTACCTGAGTCCCTA 3180
Qy 3181 CTCTACATGGGGGTCCCGAGATGAAAGAGACATGTGCTTTTATTAATTAATTAATTTGG 3240
Db 3181 CTCTACATGGGGGTCCCGAGATGAAAGAGACATGTGCTTTTATTAATTAATTAATTTGG 3240
Qy 3241 TGGTCTGTGTTATTAAGAGATCAAAATGTATTAACAACATAGCTCTTTTCACTGACTTA 3300
Db 3241 TGGTCTGTGTTATTAAGAGATCAAAATGTATTAACAACATAGCTCTTTTCACTGACTTA 3300
Qy 3301 GTATTAATCACTAATCACTAGTTTGAATGCTGGGTGTGACTTCTACTGACCGTATAGT 3360
Db 3301 GTATTAATCACTAATCACTAGTTTGAATGCTGGGTGTGACTTCTACTGACCGTATAGT 3360
Qy 3361 AAAGGTGCTGTGCTCCAGGTGGGATTAATTAATTAATTAATTAATTAATTAATTAATTA 3420
Db 3361 AAAGGTGCTGTGCTCCAGGTGGGATTAATTAATTAATTAATTAATTAATTAATTAATTA 3420
Qy 3421 ATGTGTGTGTTGAGAGCATTTGACACATATCTGCTTGTATAGAGACTTCTCTGATTTCTC 3480
Db 3421 ATGTGTGTGTTGAGAGCATTTGACACATATCTGCTTGTATAGAGACTTCTCTGATTTCTC 3480
Qy 3481 TAGGTGCTGTGCTTATCCCATTTGTGAAATTTCAATTCCTAATCCCATTTGTCTATAGT 3540
Db 3481 TAGGTGCTGTGCTTATCCCATTTGTGAAATTTCAATTCCTAATCCCATTTGTCTATAGT 3540
Qy 3541 CCTAGCAATTAAGAAATTTCTCAAGTTTCCATGTGCGGTCTCTTACGTGAGCAATTA 3600
Db 3541 CCTAGCAATTAAGAAATTTCTCAAGTTTCCATGTGCGGTCTCTTACGTGAGCAATTA 3600
Qy 3601 CTTTGACATTTAAAGAGAAATTTAGAGATATTTCTCATCTCTTAAATATTAATTAATTA 3660
Db 3601 CTTTGACATTTAAAGAGAAATTTAGAGATATTTCTCATCTCTTAAATATTAATTAATTA 3660
Qy 3661 TACCAAAAGTGGCCCCCTGATTAATTTCTGTGCGCACTGCAACCATTAATCTGATAG 3720
Db 3661 TACCAAAAGTGGCCCCCTGATTAATTTCTGTGCGCACTGCAACCATTAATCTGATAG 3720
Qy 3721 CTTTAAAAACAACATTAAGCTTATAGTCTGTGGGATCAAGATTCGAAATGGAATGCTCCT 3780
Db 3721 CTTTAAAAACAACATTAAGCTTATAGTCTGTGGGATCAAGATTCGAAATGGAATGCTCCT 3780
Qy 3781 GAATGAATTAAGAGGTGTCAGACAGAGCTGTGCTCTCTTGAAGGCTCTAGGAGAGAGCG 3840
Db 3781 GAATGAATTAAGAGGTGTCAGACAGAGCTGTGCTCTCTTGAAGGCTCTAGGAGAGAGCG 3840
Qy 3841 GTTCTTGCATTTCAAGCTTTCTAGAGGCTGTGATTCCTCAGGCTCAGTGTGCTGTC 3900
Db 3841 GTTCTTGCATTTCAAGCTTTCTAGAGGCTGTGATTCCTCAGGCTCAGTGTGCTGTC 3900
Qy 3901 AAGCTTTTCTCAATGGAATCACTGTGACACTGTGCTCTCTCTCTCTCTCTCTCTCTCTCT 3960
Db 3901 AAGCTTTTCTCAATGGAATCACTGTGACACTGTGCTCTCTCTCTCTCTCTCTCTCTCTCT 3960
Qy 3961 AAAGCCCAAGAGAAATTCAGAGATTAATCTCTCATCTTAAAGATCTTTCATCATCTGGA 4020
Db 3961 AAAGCCCAAGAGAAATTCAGAGATTAATCTCTCATCTTAAAGATCTTTCATCATCTGGA 4020
Qy 4021 AGAGCCTTTTGCATGCAAGACAAATTAAGCAAGGTGGGATTAAGACAGAGACATTT 4080
Db 4021 AGAGCCTTTTGCATGCAAGACAAATTAAGCAAGGTGGGATTAAGACAGAGACATTT 4080
Qy 4081 TGGGTGTGTTATTTCTGCTTACCAACACTTCTGCACTGACCTCCACAGAGAGGCTA 4140
Db 4081 TGGGTGTGTTATTTCTGCTTACCAACACTTCTGCACTGACCTCCACAGAGAGGCTA 4140
Qy 4141 CAAATATGATCTGGGCAAGAGATTTTGTATGCTTGGGAGCTTAACTTAAATAA 4200
Db 4141 CAAATATGATCTGGGCAAGAGATTTTGTATGCTTGGGAGCTTAACTTAAATAA 4200
Qy 4201 ACCCAATCAAGAAATCTGGCAATGTGGGCTCAATTTCTCACTTGAACAACTGGC 4260
Db 4201 ACCCAATCAAGAAATCTGGCAATGTGGGCTCAATTTCTCACTTGAACAACTGGC 4260
```

OY	4261	TGGAGCTGGGACACAGCTCTGCTTTTAAAGGGGGTGCACCTTACACGAGTCAACAAGC	4320
Dp	4261	TGGAGCTGGGACACAGCTCTGCTTTTAAAGGGGGTGCACCTTACACGAGTCAACAAGC	4320
OY	4321	CCACACTACGCCCTATCACTTCCCAACATGAGGCTAAGTGTGTTTCTACTGATCATG	4380
Dp	4321	CCACACTACGCCCTATCACTTCCCAACATGAGGCTAAGTGTGTTTCTACTGATCATG	4380
OY	4381	CCCTGCGAGGTTCATTTATTTGTAATGAAAAAGAAAGCTGGGATTAATCTCTAATCAGG	4440
Dp	4381	CCCTGCGAGGTTCATTTATTTGTAATGAAAAAGAAAGCTGGGATTAATCTCTAATCAGG	4440
OY	4441	TGAGTAGACCATGAGACCAATGTGTGGCTCACTTAACCTTTTCTTTTCTTTTCT	4500
Dp	4441	TGAGTAGACCATGAGACCAATGTGTGGCTCACTTAACCTTTTCTTTTCTTTTCT	4500
OY	4501	TTTTCTTTTTTTTTTAAATGTAGACAGAGATCTCATTTCTGTGTGCCTAGGCTGAGTGCAG	4560
Dp	4501	TTTTCTTTTTTTTTTAAATGTAGACAGAGATCTCATTTCTGTGTGCCTAGGCTGAGTGCAG	4560
OY	4561	TGGCGCAATCTGCGCTCACTGCAACTCTGTGCTCTGCGGCTCAAGCAATCTCCCACTC	4620
Dp	4561	TGGCGCAATCTGCGCTCACTGCAACTCTGTGCTCTGCGGCTCAAGCAATCTCCCACTC	4620
OY	4621	AGCTCCCAATATAGCTGGGATCATCTGGGACAAACACATGCTCCAGCTAAATTTGTATTT	4680
Dp	4621	AGCTCCCAATATAGCTGGGATCATCTGGGACAAACACATGCTCCAGCTAAATTTGTATTT	4680
OY	4681	TTTGTAGAGACAGGGTTTCAACCATTTGCCAGGCTGTCTCAACTCTCTGGGCTCAAGC	4740
Dp	4681	TTTGTAGAGACAGGGTTTCAACCATTTGCCAGGCTGTCTCAACTCTCTGGGCTCAAGC	4740
OY	4741	AATCTCTGCTCTGCGCTCTGCGCTCCCAAGTCTGGGATTAACAGATGTAGGCGACCGATCCAG	4800
Dp	4741	AATCTCTGCTCTGCGCTCTGCGCTCCCAAGTCTGGGATTAACAGATGTAGGCGACCGATCCAG	4800
OY	4801	CCCCACACCTTCATTTATTAACCAATTACTGCGCCAGTAATCTGAGACTTTTGCTTCTAC	4860
Dp	4801	CCCCACACCTTCATTTATTAACCAATTACTGCGCCAGTAATCTGAGACTTTTGCTTCTAC	4860
OY	4861	CCCTGCTCTGATCTGGAAGGAGAGGGATTAATGTTTACTGTTGTGCGACAGTCCCAAGTT	4920
Dp	4861	CCCTGCTCTGATCTGGAAGGAGAGGGATTAATGTTTACTGTTGTGCGACAGTCCCAAGTT	4920
OY	4921	CAATATTTCTGGCGCAAAAACCTTCTTCAAAAAATAAATGTAATTCATTTGATTAATGA	4980
Dp	4921	CAATATTTCTGGCGCAAAAACCTTCTTCAAAAAATAAATGTAATTCATTTGATTAATGA	4980
OY	4981	ATTCACTCTTGAAGATGCAACCGCTTAACTTTGTCACTGCGCATTAATGAAGAAATTTTA	5040
Dp	4981	ATTCACTCTTGAAGATGCAACCGCTTAACTTTGTCACTGCGCATTAATGAAGAAATTTTA	5040
OY	5041	TAGTCTCTAAATGGCGTGTACTGCAAGACCTTTGAACACTTTTCTCAAGAGATGAGATAT	5100
Dp	5041	TAGTCTCTAAATGGCGTGTACTGCAAGACCTTTTGAACACTTTTCTCAAGAGATGAGATAT	5100
OY	5101	TTAAGTCAATGCCCTTGGCGTTGCTATGAGCACTTTCCCTCTGAAATCTGTGTTCTCTGC	5160
Dp	5101	TTAAGTCAATGCCCTTGGCGTTGCTATGAGCACTTTCCCTCTGAAATCTGTGTTCTCTGC	5160
OY	5161	CCAAGTACCTTTGGCTTGTGTAGCGCGAATGCTGACCTTCATTAAGGGCCAAAGAGAGG	5220
Dp	5161	CCAAGTACCTTTGGCTTGTGTAGCGCGAATGCTGACCTTCATTAAGGGCCAAAGAGAGG	5220
OY	5221	CTGCGGCTTCTCTTCCCTCACTGAAAGGCTTATTTGAATTCACTGTGTGAGCCCTTAGC	5280
Dp	5221	CTGCGGCTTCTCTTCCCTCACTGAAAGGCTTATTTGAATTCACTGTGTGAGCCCTTAGC	5280
OY	5281	CCTCAATTTCTGCAACATCCCAACCTCCAGCCCTTCOAAGCAGAGCATAGGTGCCAG	5340
Dp	5281	CCTCAATTTCTGCAACATCCCAACCTCCAGCCCTTCOAAGCAGAGCATAGGTGCCAG	5340

QY	5341	ATTCCACCCCAAGGTGGGATTTGGCCCTTCTTAAGGCTGGCTACTTGTCAACATCACCGACAT	5400
Db	5341	ATTCCACCCCAAGGTGGGATTTGGCCCTTCTTAAGGCTGGCTACTTGTCAACATCACCGACAT	5400
QY	5401	CACGTGTCCTGCAAGGACCAACGTCGGCCATTTTCTTCACTGAGGGCTCAAACTCC	5460
Db	5401	CACGTGTCCTGCAAGGACCAACGTCGGCCATTTTCTTCACTGAGGGCTCAAACTCC	5460
QY	5461	TGGAACAAGTTGTGGCTCCGAGACAGATATTTCTGTGAAGCTGTGCTCAGTGAAGGGGC	5520
Db	5461	TGGAACAAGTTGTGGCTCCGAGACAGATATTTCTGTGAAGCTGTGCTCAGTGAAGGGGC	5520
QY	5521	CCAGCCGAGGAACCCCTGGCTCTTTTCTTTAAAGCCAGGAGCCCACTTACATAAACATT	5580
Db	5521	CCAGCCGAGGAACCCCTGGCTCTTTTCTTTAAAGCCAGGAGCCCACTTACATAAACATT	5580
QY	5581	TCAGGGTCACTGGAAACAGTGAAGTGCATTTGTGAAGCTACTGATGCCAGCCACT	5640
Db	5581	TCAGGGTCACTGGAAACAGTGAAGTGCATTTGTGAAGCTACTGATGCCAGCCACT	5640
QY	5641	GCTCATCCACGTGTCGTGCAATGCTTACGAGGAAGGCCACGCCATGACAGACTGTCTCT	5700
Db	5641	GCTCATCCACGTGTCGTGCAATGCTTACGAGGAAGGCCACGCCATGACAGACTGTCTCT	5700
QY	5701	AATGCTGTGGTCATTGTCAGAGAAAGGAAAGGTCCTCAAGAAAGTCAACTGGGACAAACA	5760
Db	5701	AATGCTGTGGTCATTGTCAGAGAAAGGAAAGGTCCTCAAGAAAGTCAACTGGGACAAACA	5760
QY	5761	CAAGCCACCGGACATGCGCTTGGTAAAGGTTAGCAGACTGATGTGTGTGGATCTGCAGT	5820
Db	5761	CAAGCCACCGGACATGCGCTTGGTAAAGGTTAGCAGACTGATGTGTGTGGATCTGCAGT	5820
QY	5821	GCTTCACTGGAATATTTATTTCAATTGCAATCTTTTAAAGTGGCAATTTATTCATTTTC	5880
Db	5821	GCTTCACTGGAATATTTATTTCAATTGCAATCTTTTAAAGTGGCAATTTATTCATTTTC	5880
QY	5881	CTGGGCTTTAAATPAAACAAATGTACCAAAAAACAAGTACAAGCTGTTAAAGTCTCCGG	5940
Db	5881	CTGGGCTTTAAATPAAACAAATGTACCAAAAAACAAGTACAAGCTGTTAAAGTCTCCGG	5940
QY	5941	CTACTTTCCTCCCTGGTTCAGTAGAGGCCCGGTTTCCAGTTGTTGACTGTGACAGGCTC	6000
Db	5941	CTACTTTCCTCCCTGGTTCAGTAGAGGCCCGGTTTCCAGTTGTTGACTGTGACAGGCTC	6000
QY	6001	AGCATGGGCTCAGCAGATGCTGCTTAATTTGTGGAATACAGAAAGCCAGGCTTTGGG	6060
Db	6001	AGCATGGGCTCAGCAGATGCTGCTTAATTTGTGGAATACAGAAAGCCAGGCTTTGGG	6060
QY	6061	ATACAAATTCCTTCCTCTTCAATTTGAAGCCGTGACGTGATGTAAGAAAGCATGTTTTCGC	6120
Db	6061	ATACAAATTCCTTCCTCTTCAATTTGAAGCCGTGACGTGATGTAAGAAAGCATGTTTTCGC	6120
QY	6121	GGAATATAAATAATAGTCTTGGAGTCTCCGCCAAAAAATAAAAA 6163	
Db	6121	GGAATATAAATAATAGTCTTGGAGTCTCCGCCAAAAAATAAAAA 6163	
RESULT 2			
US-09-814-353-20489			
Sequence 20489, Application US/09814353			
Publication No. US20030165831A1			
GENERAL INFORMATION:			
APPLICANT: Lee, John			
APPLICANT: Thompson, Pamela			
APPLICANT: Lillie, James			
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR			
IDENTIFICATION, ASSESSMENT, PREVENTION, AND			
THERAPY OF OVARIAN CANCER			
FILE REFERENCE: MRI-006B			
CURRENT APPLICATION NUMBER: US/09/814,353			
CURRENT FILING DATE: 2001-03-21			
PRIOR APPLICATION NUMBER: US 60/191,031			
PRIOR FILING DATE: 2000-03-21			

PRIOR APPLICATION NUMBER: US 60/207,124
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 60/211,940
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: US 60/216,820
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: US 60/220,661
PRIOR FILING DATE: 2000-07-25
PRIOR APPLICATION NUMBER: US 60/257,672
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 22037
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 20489
LENGTH: 6286
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1, 2, 3, 4, 5, 6284, 6285, 6286
OTHER INFORMATION: n = A,T,C or G
US-09-814-353-20489

Query Match 87.6%; Score 5400; DB 13; Length 6286;
Beet Local Similarity 99.9%; Pred. No. 0;
Matches 5840; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

309 TATCCAGCTAGTGGGTTCTGCTGCTGGGCGCGCTGCGCGGGGCGAGAAC 368
66 TATCCAGCTAGTGGGTTCTGCTGCTGGGCGCGCTGCGCGGGGCGAGAAC 125
369 TTTTGAAGTTGCTGCGCAGAGAAAGCAATTCAGTTCTATTAAGCTGGGAGCCC 428
126 TTTTGAAGTTGCTGCGCAGAGAAAGCAATTCAGTTCTATTAAGCTGGGAGCCC 185
429 GACTGCTGCGCAAAACCTGTTACATGCTCATTTTAAAGACATATTAACATGTTGTC 488
186 GACTGCTGCGCAAAACCTGTTACATGCTCATTTTAAAGACATATTAACATGTTGTC 245
489 CATCAAGCTGAGAGAAAGATAGTCTTACCTTTAGCTGCGAGAGCTGAGATCAGTT 548
246 CATCAAGCTGAGAGAAAGATAGTCTTACCTTTAGCTGCGAGAGCTGAGATCAGTT 305
549 TGTCTAGAGATCCAGAAAAATATTAAGCTGATATGTCAGGCCCATGTCCTTTGGGAGGT 608
306 TGTCTAGAGATCCAGAAAAATATTAAGCTGATATGTCAGGCCCATGTCCTTTGGGAGGT 365
609 TAGCTTACGCTTGAATCTGTTGCTTACCTTCAACAGAACTTTTATCTGGAGTGT 668
366 TCAGCTTACGCTTGAATCTGTTGCTTACCTTCAACAGAACTTTTATCTGGAGTGT 425
669 CAAGCTCATAGAGATGAGTGTAGAGCTGAGATTTTCCATCCCTCGCTGAGGAGAT 728
426 CAAGCTCATAGAGATGAGTGTAGAGCTGAGATTTTCCATCCCTCGCTGAGGAGAT 484
729 CGGTCCGGTGAAGCTGCGCAGAGAGTCACTCATCTCAGCGGCCGATGATGTC 788
485 CGGTCCGGTGAAGCTGCGCAGAGAGTCACTCATCTCAGCGGCCGATGATGTC 544
789 CACCGTGTGAGAGTGGAACTTCTGCGAGCAATGGCACTGTGTCCTGGATCAAGATGCA 848
545 CACCGTGTGAGAGTGGAACTTCTGCGAGCAATGGCACTGTGTCCTGGATCAAGATGCA 604
849 AGAAGAGTGAAGATGGCTTACACTCCAGTGTTCACCCCAAGAAATGTCCTGGGCTT 908
605 AGAAGAGTGAAGATGGCTTACACTCCAGTGTTCACCCCAAGAAATGTCCTGGGCTT 664
909 CAGCATTTGAAACCGCTCATCTATAAAAGTCTGTCATCATGATGATGTTTGAAGG 968
665 CAGCATTTGAAACCGCTCATCTATAAAAGTCTGTCATCATGATGATGTTTGAAGG 724
969 TGAAGGCTCAGCAACCTGATGTCGCAAGTACCCAGAGGCTTCCCTGAGAGTACCT 1028
725 TGAAGGCTCAGCAACCTGATGTCGCAAGTACCCAGAGGCTTCCCTGAGAGTACCT 784

1029 CATGACGTGAGAGTTGTCGTTCTGCAACACTGCGGCGAGCGTCTCTCTCAACTT 1088
785 CATGACGTGAGAGTTGTCGTTCTGCAACACTGCGGCGAGCGTCTCTCTCAACTT 844
1089 CAACCTTCTCAACTGTGAGAGAGAGAGCGGTTGATATCTATCATCTCCGGGCTCCAC 1148
845 CAACCTTCTCAACTGTGAGAGAGAGAGCGGTTGATATCTATCATCTCCGGGCTCCAC 904
1149 CACCAACCCGAGAGGTTCAAGCTGAGAGCAAGAGCTGGGAATATGCGGGGAACCTT 1208
905 CACCAACCCGAGAGGTTCAAGCTGAGAGCAAGAGCTGGGAATATGCGGGGAACCTT 964
1209 CAACCTTCTCTGCAAGGCTGAGCAAGAGTCCCAAGATCCCAAGGATCTCCGGCTGCA 1268
965 CAACCTTCTCTGCAAGGCTGAGCAAGAGTCCCAAGATCTCCGGGATCTCCGGCTGCA 1024
1269 GTTCCAGTTTGTGTCACATCCACAAATGAAGCAATAAATCTACGTGTTGACTT 1328
1025 GTTCCAGTTTGTGTCACATCCACAAATGAAGCAATAAATCTACGTGTTGACTT 1084
1329 GAGTAATGAGAGGCAATGTCATCAGCATGAGCCAGCGCCCTTCAACAGAGCGGCA 1388
1085 GAGTAATGAGAGGCAATGTCATCAGCATGAGCCAGCGCCCTTCAACAGAGCGGCA 1144
1389 GTTGTCCCTGAGCTGTTTGTGTCATGAATCTGAGACCTGAGTACAGCACTCAGCCT 1448
1145 GTTGTCCCTGAGCTGTTTGTGTCATGAATCTGAGACCTGAGTACAGCACTCAGCCT 1204
1449 GACATCTGAGCTTCAACAAATATCTCTTGTGATGATCTGACAGCTCTGTGAT 1508
1205 GACATCTGAGCTTCAACAAATATCTCTTGTGATGATCTGACAGCTCTGTGAT 1264
1509 GAATGTGAAAAAACATTAAGCTGACAGACCGGATCTGCAAGAGAAATCTTACTC 1568
1265 GAATGTGAAAAAACATTAAGCTGACAGACCGGATCTGCAAGAGAAATCTTACTC 1324
1569 ACTCAGGTCCTGAGTGAATCTCAGCTGCTGAGAGCTGAGTGAATCTCTCTGGA 1628
1325 ACTCAGGTCCTGAGTGAATCTCAGCTGCTGAGAGCTGAGTGAATCTCTCTGGA 1384
1629 GCTGTGTCGCCAAGAGAGAGCTGAGCTGAGTGTGTCAGAGCCAGAGCTGACGA 1688
1385 GCTGTGTCGCCAAGAGAGAGCTGAGCTGAGTGTGTCAGAGCCAGAGCTGACGA 1444
1689 GCATCAACAGAGAGAGCTGAGCTGAGTGTGTCAGAGCCAGAGCTGACGA 1748
1445 GCATCAACAGAGAGAGCTGAGCTGAGTGTGTCAGAGCCAGAGCTGACGA 1504
1749 CAGCAGAGAGCTGATCTGAGCTGCTTCTGAGCCGGAGAGCTTATCAAGCAGATCAGGT 1808
1505 CAGCAGAGAGCTGATCTGAGCTGCTTCTGAGCCGGAGAGCTTATCAAGCAGATCAGGT 1564
1809 GAAGCAGAACTTCTGAGTACCTTCTGACCTTTGCGACCTTTCACAAAGAGGCTC 1868
1565 GAAGCAGAACTTCTGAGTACCTTCTGACCTTTGCGACCTTTCACAAAGAGGCTC 1624
1869 CAGCAGAGTCTGAGCTGCTTCTTATTAACCTTATTAAGAGAGAGGCTTTTCAAGT 1928
1625 CAGCAGAGTCTGAGCTGCTTCTTATTAACCTTATTAAGAGAGAGGCTTTTCAAGT 1684
1929 GACCCCTGACCAAAAGAGAGTCTACCTGAGAGCCCAAGTGGAGCCGGGAGCTGCC 1988
1685 GACCCCTGACCAAAAGAGAGTCTACCTGAGAGCCCAAGTGGAGCCGGGAGCTGCC 1744
1989 ATCCCTCACTCTGTGTCTGTAACATCAGAGTGCACAGAGCCAGTGTGCTGCTAC 2048
1745 ATCCCTCACTCTGTGTCTGTAACATCAGAGTGCACAGAGCCAGTGTGCTGCTAC 1804
2049 TTTCTTTAAGAGCGGAGCGGCTGTGTCAGACAGAGGCGCGATTCATGATCATCA 2108
1805 TTTCTTTAAGAGCGGAGCGGCTGTGTCAGACAGAGGCGCGATTCATGATCATCA 1864

2109 GGAGCAGCGGACCCGGGCTGAGAGATCTTCAAGCTGAGCAGAGATGTCTCCCAAGCC 2168
1865 GGAGCAGCGGACCCGGGCTGAGAGATCTTCAAGCTGAGCAGAGATGTCTCCCAAGCC 1924
2169 AAGCTTCCAACATCAGAGCTTGGGTCACATCTCTTAATGAGGCTCCAGAGCGGCAA 2228
1925 AAGCTTCCAACATCAGAGCTTGGGTCACATCTCTTAATGAGGCTCCAGAGCGGCAA 1984
2229 GAGCTAGACCTGCTCTTCCGTGACACTTACCCCAAGGACCTGAGACTTGACCTGTCA 2288
1985 GAGCTAGACCTGCTCTTCCGTGACACTTACCCCAAGGACCTGAGACTTGACCTGTCA 2044
2289 CCTCATGACGCGGTGGAGGTGAGTCTTACTGCTGTCTGCTCGGCTCATCATTTTG 2348
2045 CCTCATGACGCGGTGGAGGTGAGTCTTACTGCTGTCTGCTCGGCTCATCATTTTG 2104
2349 CTGTGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 2408
2105 CTGTGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 2164
2409 CAACATCAATACTGAGATGCCGAGGAGCCAAAGAAAGTTTCAGAAAGGCGAAAGACAA 2468
2165 CAACATCAATACTGAGATGCCGAGGAGCCAAAGAAAGTTTCAGAAAGGCGAAAGACAA 2224
2469 TGACTCCCATGTGTATGACATCATCGAGGACACATGTATATGGCATCTGTACAGGA 2528
2225 TGACTCCCATGTGTATGACATCATCGAGGACACATGTATATGGCATCTGTACAGGA 2284
2529 TTCCAGCGGCTCTTCTGTGAGCCAGAGGTGACACCTTACCGGCGTTCCAGGCGCCAT 2588
2285 TTCCAGCGGCTCTTCTGTGAGCCAGAGGTGACACCTTACCGGCGTTCCAGGCGCCAT 2344
2589 GGGGGTCTGTCTCTCCCTCCCAACCAACATATGCTCAAGGCGCCCACTGCAAAAGTTGGC 2648
2345 GGGGGTCTGTCTCTCCCTCCCAACCAACATATGCTCAAGGCGCCCACTGCAAAAGTTGGC 2404
2649 CACTGAGAGCCACTTCTGTGCTCCCTCTGAGTCTGAGAGTGAACCGTACACTTCTC 2708
2405 CACTGAGAGCCACTTCTGTGCTCCCTCTGAGTCTGAGAGTGAACCGTACACTTCTC 2464
2709 CCATCCCAACATGGGGATGTATGAGCAAGGACACAGACATTCCTTCTGTGAACCTCA 2768
2465 CCATCCCAACATGGGGATGTATGAGCAAGGACACAGACATTCCTTCTGTGAACCTCA 2524
2769 GGAGCCCATGAGCCAGCAAGATTAATTCATTCAGACGCTTGTGATTTCA 2828
2525 GGAGCCCATGAGCCAGCAAGATTAATTCATTCAGACGCTTGTGATTTCA 2584
2829 AAGCAGGGCACTGAGACACCCGTCGTTCTTAACCAAGAAATCTTAAAGAGAGAAAT 2888
2585 AAGCAGGGCACTGAGACACCCGTCGTTCTTAACCAAGAAATCTTAAAGAGAGAAAT 2644
2889 ATACAGAAAGAACAGAGAGGTTTCTGAGACAGCGCAACTTCAATTTGCTCACTGGA 2948
2645 ATACAGAAAGAACAGAGAGGTTTCTGAGACAGCGCAACTTCAATTTGCTCACTGGA 2704
2949 CTCATTCTAAGGCGCAAGACTTGAATGATTAATTCATCTGATTAAGTCACTGAGCAG 3008
2705 CTCATTCTAAGGCGCAAGACTTGAATGATTAATTCATCTGATTAAGTCACTGAGCAG 2764
3009 CTCATGTGCTCTCAACTTAAGGCTGTGCGGTTAGCCAGCCTGTATATGAGAGAGAGC 3068
2765 CTCATGTGCTCTCAACTTAAGGCTGTGCGGTTAGCCAGCCTGTATATGAGAGAGAGC 2824
3069 CTGAGTCACTTACATAGGTTTGACAGAGCCCTGATTCAGAGCTTAAACAGAGCTT 3128
2825 CTGAGTCACTTACATAGGTTTGACAGAGCCCTGATTCAGAGCTTAAACAGAGCTT 2884
3129 GCCCTTTGAGAGCAACAGTTTCAATTTCAAGAGGCTTCTGAGTCTCTCACT 3188
2885 GCCCTTTGAGAGCAACAGTTTCAATTTCAAGAGGCTTCTGAGTCTCTCACT 2944
3189 GGGGTCCCAAGATGAAGAAAGCAATATGTGCTTTTATTAATTAATTTGTTGTTGCTG 3248

2945 GGGGTCCCAAGATGAAGAAAGCAATATGTGCTTTTATTAATTAATTTGTTGTTGCTG 3004
3249 TGTATTTAAGAGATCAATATGATTAACACAGCTCTTTTACCTGACTTAATAAC 3308
3005 TGTATTTAAGAGATCAATATGATTAACACAGCTCTTTTACCTGACTTAATAAC 3064
3309 TCATATCACTGAGTTTGAATGCTGGTTTGAATCTTCACTGACCGCTAGATTAACGCT 3368
3065 TCATATCACTGAGTTTGAATGCTGGTTTGAATCTTCACTGACCGCTAGATTAACGCT 3124
3369 GGGTCTCCCAAGGAGGTGGGATTAATTTAACAATGTCACCAAGAAAGAAAGTGTGT 3428
3125 GGGTCTCCCAAGGAGGTGGGATTAATTTAACAATGTCACCAAGAAAGAAAGTGTGT 3184
3429 GTTTAGCAGCAATTAACATATCTGTTTGAATTAAGACCTTCTGATTTCTAGTCCG 3488
3185 GTTTAGCAGCAATTAACATATCTGTTTGAATTAAGACCTTCTGATTTCTAGTCCG 3244
3489 TTGTTGTTATCCATTTGTGAAATTCATTTGAATCCCATTTGCTTATATGCTTACAA 3548
3245 TTGTTGTTATCCATTTGTGAAATTCATTTGAATCCCATTTGCTTATATGCTTACAA 3304
3549 TTAAGAAATTTCTCAAGTTTCCATGTCGGTTCTCTAGCTGAGCAATCTTTGA 3608
3305 TTAAGAAATTTCTCAAGTTTCCATGTCGGTTCTCTAGCTGAGCAATCTTTGA 3364
3609 TTTAAGAAATTTAGAAATATCTCATCTCTTAAATTTTAAATATATTAACAAAC 3668
3365 TTTAAGAAATTTAGAAATATCTCATCTCTTAAATTTTAAATATATTAACAAAC 3424
3669 AGTGCCCCCTGACTTAAGTTTCTGTGCACTGCAACCCATTAATCTTGTAGCTTAA 3728
3425 AGTGCCCCCTGACTTAAGTTTCTGTGCACTGCAACCCATTAATCTTGTAGCTTAA 3484
3729 CAACATTAAGTTTATGCTTGGGGATCAGAAATTCAAAGAAAGTGTGCTGAATGA 3788
3485 CAACATTAAGTTTATGCTTGGGGATCAGAAATTCAAAGAAAGTGTGCTGAATGA 3544
3789 ATCAAGTGTGAGCAGAGCTGTGCTCTTCTGAAAGCTCTAGAGAAAGCCGTTCTTG 3848
3545 ATCAAGTGTGAGCAGAGCTGTGCTCTTCTGAAAGCTCTAGAGAAAGCCGTTCTTG 3604
3849 CCATTTCAAGCTTCTAGAGGCTGTGCTGCAATCCAGGCTCCAGTGTGTCAAGCTTTT 3908
3605 CCATTTCAAGCTTCTAGAGGCTGTGCTGCAATCCAGGCTCCAGTGTGTCAAGCTTTT 3664
3909 CTCATGAGCACTGAGGACCTGGCCCTCCACTTCCCTTGTGAATTAAGAGCCCA 3968
3665 CTCATGAGCACTGAGGACCTGGCCCTCCACTTCCCTTGTGAATTAAGAGCCCA 3724
3969 CCAGAAAGTCCAGGATATCTCTCCATCTAAGATCTTCACTCATCTGAAAGAGCTT 4028
3725 CCAGAAAGTCCAGGATATCTCTCCATCTAAGATCTTCACTCATCTGAAAGAGCTT 3784
4029 TTGCAATGCAAGCAACATTAAGCAGAGGTGGGATTAAGACCAAGACATCTTTGGGTC 4088
3785 TTGCAATGCAAGCAACATTAAGCAGAGGTGGGATTAAGACCAAGACATCTTTGGGTC 3844
4089 TGTATTTCTGCTTACCAACTTCTGCACTGACCTCCACAGAGAGGCTTCAAAATGA 4148
3845 TGTATTTCTGCTTACCAACTTCTGCACTGACCTCCACAGAGAGGCTTCAAAATGA 3904
4149 TCTGAGCAGCAGAGATGTTTGTAGCTTGGGACTTAAACATTT-AAAAAACCAG 4207
3905 TCTGAGCAGCAGAGATGTTTGTAGCTTGGGACTTAAACATTTAAAAAACCAG 3964
4208 ATCAGAAAGTCTGAGCATGCTGGGCTCAATTTCTACCTAGCAACATGCTGAGCT 4267
3965 ATCAGAAAGTCTGAGCATGCTGGGCTCAATTTCTACCTAGCAACATGCTGAGCT 4024
4268 GGGCACAGCTTGTCTTTAAGAGGGGTGTCACTTACAGAGTCAACACAGCCCACT 4327

Db 4025 GGGACCAAGCTGCTGCTTTAGAAAGGGGTGTCCATTACACAGGTACCAAGCCCACT 4084
Qy 4328 AGCCCTATCACTTCCCAATAGAGGCTAAGTGTGTTTCTACTGATCAATGCCCTGC 4387
Db 4085 AGCCCTATCACTTCCCAATAGAGGCTAAGTGTGTTTCTACTGATCAATGCCCTGC 4144
Qy 4388 AGGTTCATTTATTTGTAATGAAAAAGAAAGACTGGAGTTAATCTCTAATCAGGTAGT 4447
Db 4145 AGGTTCATTTATTTGTAATGAAAAAGAAAGACTGGAGTTAATCTCTAATCAGGTAGT 4204
Qy 4448 ACCATGAGACCAATGAGTGTCACTAACCTTTTCTTTTCTTTTCTTTTCTTTTCTT 4507
Db 4205 ACCATGAGACCAATGAGTGTCACTAACCTTTTCTTTTCTTTTCTTTTCTTTTCTT 4264
Qy 4508 TTTTCTTTTAAATGAGACAGGATCTATTCTGTTGCTAGGCTGGAAGTGGGCGA 4567
Db 4265 TTTTCTTTTAAATGAGACAGGATCTATTCTGTTGCTAGGCTGGAAGTGGGCGA 4324
Qy 4568 ATCTGGCTCACTGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4627
Db 4325 ATCTGGCTCACTGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4384
Qy 4628 CAATAGCTGGAGTACATGCGCAAAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4687
Db 4385 CAATAGCTGGAGTACATGCGCAAAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4444
Qy 4688 AGACAGGCTTCAACCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4747
Db 4445 AGACAGGCTTCAACCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4504
Qy 4748 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4807
Db 4505 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4564
Qy 4808 CCTCATTTATTAACCAATTAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4867
Db 4565 CCTCATTTATTAACCAATTAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4624
Qy 4868 CTGATCTGGAAGAGAGGATTAATGTTAATGCTTGCAGACAGTCCCAAGTTCAATAT 4927
Db 4625 CTGATCTGGAAGAGAGGATTAATGTTAATGCTTGCAGACAGTCCCAAGTTCAATAT 4684
Qy 4928 TCTGCGGCAAAAACCTTCTTCAAAAATAATGTAATTGTAATGTAATGTAATGTAAT 4987
Db 4685 TCTGCGGCAAAAACCTTCTTCAAAAATAATGTAATTGTAATGTAATGTAATGTAAT 4744
Qy 4988 TTGGAATAGCACCGGCTCACTGTTGCATGCGCAATAATGAAAGAAATTTATATGCTC 5047
Db 4745 TTGGAATAGCACCGGCTCACTGTTGCATGCGCAATAATGAAAGAAATTTATATGCTC 4804
Qy 5048 CTAAATGCGGTGTACTGCAAGACCTTGAACACTTCCAGAGGATAGGATATTTAAGTC 5107
Db 4805 CTAAATGCGGTGTACTGCAAGACCTTGAACACTTCCAGAGGATAGGATATTTAAGTC 4864
Qy 5108 ATGCCCTTGGGCTTGCCTATGCGACCTTTCCTTCTGAAAGCTGCTGCTGCTGCTGCT 5167
Db 4865 ATGCCCTTGGGCTTGCCTATGCGACCTTTCCTTCTGAAAGCTGCTGCTGCTGCTGCT 4924
Qy 5168 CCTTGGCTTGTGAGCGCGAGATGCTGCACTGCAATAAGGCGCAAGAGAGGCTGCGGC 5227
Db 4925 CCTTGGCTTGTGAGCGCGAGATGCTGCACTGCAATAAGGCGCGAGAGAGGCTGCGGC 4984
Qy 5228 TTCTCTTCCCTCACTGAGAGGCTTATTTGAATTACAGTGTGAGGCTTGCCTGCTGCAT 5287
Db 4985 TTCTCTTCCCTCACTGAGAGGCTTATTTGAATTACAGTGTGAGGCTTGCCTGCTGCAT 5044
Qy 5288 TCTGACATTTCCCAACCTCCAGGCTTCCCAAGCAGAGACTAGTGGCTGCTGCTGCTGCT 5347
Db 5045 TCTGACATTTCCCAACCTCCAGGCTTCCCAAGCAGAGACTAGTGGCTGCTGCTGCTGCT 5104
Qy 5348 CCAAGGTGGATTTGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5407
Db 5105 CCAAGGTGGATTTGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5164

Qy 5408 GCTTCAAGACACACGCTGCGCATTTTCTTCAACTGAGGCTCAAAAACCTCTGACAA 5467
Db 5165 GCTTCAAGACACACGCTGCGCATTTTCTTCAACTGAGGCTCAAAAACCTCTGACAA 5224
Qy 5468 GTTGTGCTCTCTGAGACCAAGTATTTCTTGAGCTGTGCTCTGATGAAAGGGGCCAGCT 5527
Db 5225 GTTGTGCTCTCTGAGACCAAGTATTTCTTGAGCTGTGCTCTGATGAAAGGGGCCAGCT 5284
Qy 5528 GAGGAACCTGCTCTTCTTCTTAAAGCCAGGCCCCCACTTAACATAAAGCTTAAAGGT 5587
Db 5285 GAGGAACCTGCTCTTCTTCTTAAAGCCAGGCCCCCACTTAACATAAAGCTTAAAGGT 5344
Qy 5588 CACTGGAACAGTGAAGTGCATTTGTTGAAGCTTACAGCAATGCGAGCCACTGCTCATC 5647
Db 5345 CACTGGAACAGTGAAGTGCATTTGTTGAAGCTTACAGCAATGCGAGCCACTGCTCATC 5404
Qy 5648 CACGTGTCTGCTGCTGCTTACGAGAGGCGCAGCGCATGCAAGACTGCTTCTTAATGCTG 5707
Db 5405 CACGTGTCTGCTGCTGCTTACGAGAGGCGCAGCGCATGCAAGACTGCTTCTTAATGCTG 5464
Qy 5708 TGTCTATTGCAACAAAGGAAAGGCTCTCAAGAAAGTCAACTGGGACAAAGCAAGCC 5767
Db 5465 TGTCTATTGCAACAAAGGAAAGGCTCTCAAGAAAGTCAACTGGGACAAAGCAAGCC 5524
Qy 5768 ACCGACATGCGCTTGTGTAAGTTAGCAGCTGCTGTGTGATCTGCAAGTCTTGCAC 5827
Db 5525 ACCGACATGCGCTTGTGTAAGTTAGCAGCTGCTGTGTGATCTGCAAGTCTTGCAC 5584
Qy 5828 TGAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5887
Db 5585 TGAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5644
Qy 5888 TTAATAATAACAAATGTAACCAAAAACAAATGTAACAAATGTAACAAATGTAACAAAT 5947
Db 5645 TTAATAATAACAAATGTAACCAAAAACAAATGTAACAAATGTAACAAATGTAACAAAT 5704
Qy 5948 TCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6007
Db 5705 TCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5764
Qy 6008 GCTCAGCAGATGCTGCTTAAATTTGTGATGATACAGAAAGCCAGGCTTTGGATACAG 6067
Db 5765 GCTCAGCAGATGCTGCTTAAATTTGTGATGATACAGAAAGCCAGGCTTTGGATACAG 5824
Qy 6068 TTTCTTCTCTTCAATTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6127
Db 5825 TTTCTTCTCTTCAATTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5884
Qy 6128 AAAATTAATGCTTGGAGTCTGCGCA 6153
Db 5885 AAAATTAATGCTTGGAGTCTGCGCA 5910

RESULT 3
US-09-899-569a-1
; Sequence 1, Application US/0989569A
; Patent No. US20020142003A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020142003Albert Schweifer
; APPLICANT: Marwa Scherl-Mostageer
; APPLICANT: Wolfgang Sommergruber
; APPLICANT: Roger Abscher
; TITLE OF INVENTION: Tumorasozietres Antigen (B345)
; FILE REFERENCE: 0652.2280001
; CURRENT FILING DATE: US/09/899, 569A
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: DE 100 33 080.0
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: DE 101 19 294.0
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 60/243,158
; PRIOR FILING DATE: 2000-10-25

PRIOR APPLICATION NUMBER: US 60/297,747
PRIOR FILING DATE: 2001-06-14
NUMBER OF SEQ ID NOS: 40
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO. 1
LENGTH: 5897
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (5348)...(5348)
OTHER INFORMATION: n is any nucleotide of a, t, g or c
NAME/KEY: misc_feature
LOCATION: (5360)...(5360)
OTHER INFORMATION: n is any nucleotide of a, t, g or c
NAME/KEY: misc_feature
LOCATION: (5425)...(5425)
OTHER INFORMATION: n is any nucleotide of a, t, g or c
NAME/KEY: misc_feature
LOCATION: (5438)...(5438)
OTHER INFORMATION: n is any nucleotide of a, t, g or c
NAME/KEY: 5' UTR
LOCATION: (1)...(214)
NAME/KEY: CDS
LOCATION: (215)...(2464)
NAME/KEY: 3' UTR
LOCATION: (2465)...(5897)
US-09-899-569A-1

Query Match 87.5%; Score 5395; DB 10; Length 5897;

Best Local Similarity 99.9%; Pred. No. 0; Mismatches 7; Indels 1; Gaps 1;

Matches 5865; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

291 CCTGAATCGGGGCTCTATCGCACTGCTAGGGGTTCTGCTGGTGGCGGCGCT 350
25 CCTGAATCGGGGCTCTATCGCACTGCTAGGGGTTCTGCTGGTGGCGGCGCT 84
351 GCGCGCGGGGCGAAGCTTTGAGATTGCTCTGCCAGAGAAAGCAATTACATTCT 410
85 GCGCGCGGGGCGAAGCTTTGAGATTGCTCTGCCAGAGAAAGCAATTACATTCT 144
411 CATTAAGCTGGGAGACCCGACCTGCTGGCAAAACCTGTACATGTCATTTCTTAAAG 470
145 CATTAAGCTGGGAGACCCGACCTGCTGGCAAAACCTGTACATGTCATTTCTTAAAG 204
471 ACATATAACCATGTTTCATCAAGTCTGGAGAAAGATAGTCTTACCTTTAGCTGCA 530
205 ACATATAACCATGTTTCATCAAGTCTGGAGAAAGATAGTCTTACCTTTAGCTGCA 264
531 GAGTCTGAGAAATCACTTTGTCTATAGATCCAGAAAAATATTGACTGTATGTCAAGCC 590
265 GAGTCTGAGAAATCACTTTGTCTATAGATCCAGAAAAATATTGACTGTATGTCAAGCC 324
591 ATGTCTTTTGGGAGAGTTCAAGTCTGAGCCCTGCAATCGTTTGGCTTACCTCAAG 650
325 ATGTCTTTTGGGAGAGTTCAAGTCTGAGCCCTGCAATCGTTTGGCTTACCTCAAG 384
651 AACTTTCATCTGGAGATGTCATAAGCTCATTAAGACATCGTTTGAAGCTCAAGTTTTCAT 710
385 AACTTTCATCTGGAGATGTCATAAGCTCATTAAGACATCGTTTGAAGCTCAAGTTTTCAT 444
711 CCTCTGCTGAGGCAAGTCGGTCCGGGTGAGAGCTGCCAGAGAGTCACTCACTCAT 770
445 CCTCTGCTGAGGCAAGTCGGTCCGGGTGAGAGCTGCCAGAGAGTCACTCACTCAT 504
771 CAGCGGCGGATGATGATGCCACCGTGTCAAGATCGGAACCTTTTGCAGCAATGGCACTGT 830
505 CAGCGGCGGATGATGATGCCACCGTGTCAAGATCGGAACCTTTTGCAGCAATGGCACTGT 564
831 GTCCGGGATCAAGATGCAAGAGAGGATGAATGGCTTACACCTCCATGGTTCCAGCC 890
565 GTCCGGGATCAAGATGCAAGAGAGGATGAATGGCTTACACCTCCATGGTTCCAGCC 624

891 CAGAAATGTCCTCCGGCTTACGATTTGCAACCGCTCATCTTAAAGCTGTGTGATCAT 950
625 CAGAAATGTCCTCCGGCTTACGATTTGCAACCGCTCATCTTAAAGCTGTGTGATCAT 684
951 CGAGTCTGTGTTGAGGGTGAAGGCTCAGCAACCTGTATGTCTGCCAATCAACCGAAGG 1010
685 CGAGTCTGTGTTGAGGGTGAAGGCTCAGCAACCTGTATGTCTGCCAATCAACCGAAGG 744
1011 CTTCCCTGAGATGAGCTCATGACGTGCAAGTTTGTGTTCTCTGCAACCTGCGGCGCAG 1070
745 CTTCCCTGAGATGAGCTCATGACGTGCAAGTTTGTGTTCTCTGCAACCTGCGGCGCAG 804
1071 CTTCCCTGAGATGAGCTCATGACGTGCAAGTTTGTGTTCTCTGCAACCTGCGGCGCAG 1130
805 CTTCCCTGAGATGAGCTCATGACGTGCAAGTTTGTGTTCTCTGCAACCTGCGGCGCAG 864
1131 CTACATCCCGGGCTCAGCAACCAACCGGAGGTTTCAAGCTGAGGAGCAAGCAGCTGG 1190
865 CTACATCCCGGGCTCAGCAACCAACCGGAGGTTTCAAGCTGAGGAGCAAGCAGCTGG 924
1191 GAACATGCGGGGAACTTCAACCTCTCTCTGCAAGGCTGTGACCAAGATGCCCAAGTCC 1250
925 GAACATGCGGGGAACTTCAACCTCTCTCTGCAAGGCTGTGACCAAGATGCCCAAGTCC 984
1251 AGGATCTCTCGGCTGAGCTTCCAAAGTTTGGTCCCAATCCAGAAATGAAGCAATTA 1310
985 AGGATCTCTCGGCTGAGCTTCCAAAGTTTGGTCCCAATCCAGAAATGAAGCAATTA 1044
1311 AATCTAGCTGTTGACTGAGTAATGAGAGGCACTGCACTCACTGAGCCAGCC 1370
1045 AATCTAGCTGTTGACTGAGTAATGAGAGGCACTGCACTCACTGAGCCAGCC 1104
1371 CGTCAACAGAGCGGCAAGTTTGTCTCTGCTGTTTCTGTGTCTTAAATCTTGGACCTG 1430
1105 CGTCAACAGAGCGGCAAGTTTGTCTCTGCTGTTTCTGTGTCTTAAATCTTGGACCTG 1164
1431 CAGTAGCAACCTCAACCTGCAATCTGCTCCAAACAAAGAAATCTCTCTTGTGATGA 1490
1165 CAGTAGCAACCTCAACCTGCAATCTGCTCCAAACAAAGAAATCTCTCTTGTGATGA 1224
1491 TCTGACAGCTGTGAGTAATGTGGAAGAAACATTAAGCTGCAAGAGCAACCGGATCTG 1550
1225 TCTGACAGCTGTGAGTAATGTGGAAGAAACATTAAGCTGCAAGAGCAACCGGATCTG 1284
1551 CCAGAGAAATCTTCACTCACTGCAAGTGGCCAGTGAATCTTCACTGCTGTGAGCT 1610
1285 CCAGAGAAATCTTCACTCACTGCAAGTGGCCAGTGAATCTTCACTGCTGTGAGCT 1344
1611 GCATGACTTCTCTGGAAGCTGCTGTGCTCCAGAGACAGGCTCAAGCTGTGTGCTG 1670
1345 GCATGACTTCTCTGGAAGCTGCTGTGCTCCAGAGACAGGCTCAAGCTGTGTGCTG 1404
1671 AGCCAGAGGCTGAGAGCATACACAGAGAGGCTTGCAGACACAGCTTCAAGTCACT 1730
1405 AGCCAGAGGCTGAGAGCATACACAGAGAGGCTTGCAGACACAGCTTCAAGTCACT 1464
1731 CGTGGCCAGTGGCATACCAAGCAGAGCTGTGATCTTCTGCTCTGCTGCGGAGGCTC 1790
1465 CGTGGCCAGTGGCATACCAAGCAGAGCTGTGATCTTCTGCTCTGCTGCGGAGGCTC 1524
1791 TATCAAGAGATCAGGTTGAGAGCAAAATCTCGGTGACCTTGGCACTTTGGCCCCAG 1850
1525 TATCAAGAGATCAGGTTGAGAGCAAAATCTCGGTGACCTTGGCACTTTGGCCCCAG 1584
1851 CTTCAACAGAGGCTCAGGCAAGGCTGAGCGGTCTGATACCTTATTTCAAGAA 1910
1585 CTTCAACAGAGGCTCAGGCAAGGCTGAGCGGTCTGATACCTTATTTCAAGAA 1644
1911 GGAAGGCGTTTTCAGGTTGACCTTGAACAACAAAGGCTTCACTGAGGAGACCCCA 1970
1645 GGAAGGCGTTTTCAGGTTGACCTTGAACAACAAAGGCTTCACTGAGGAGACCCCA 1704
1971 CTGGAGCCGGGGCTGTCATCTCTGATCTGTGTCCTGGAACATGACGCTCCAGAGA 2030

Db	1705	CTGGGACCGGGGCTGGCCATCCCTCACTGTGTCCTGGAAATCAGAGTCCCAAGAA	1764
Qy	2031	CCAGGTGGCTGCTGACTTTCTTTAAAGAGCGGAGCGTGTCTGCGACAGGGCG	2090
Db	1765	CCAGGTGGCTGCTGACTTTCTTTAAAGAGCGGAGCGTGTCTGCGACAGGGCG	1824
Qy	2091	CGCATTCATGATCATCCAGAGCAGCGGACCCGGGCTGAGAGATCTTACGCTGGACA	2150
Db	1825	CCCATTCATGATCATCCAGAGCAGCGGACCCGGGCTGAGAGATCTTACGCTGGACA	1884
Qy	2151	GGATGTGCTCCCAAGCCAAAGCTTCCACATCAAGCTTCTGGGTCAATCTTAACTG	2210
Db	1885	GGATGTGCTCCCAAGCCAAAGCTTCCACATCAAGCTTCTGGGTCAATCTTAACTG	1944
Qy	2211	CAGCCCAAGAGCGGCAAGCAGCTAGACCTGCTCTTCTGGTGAACATTAACCCAAAGAC	2270
Db	1945	CAGCCCAAGAGCGGCAAGCAGCTAGACCTGCTCTTCTGGTGAACATTAACCCAAAGAC	2004
Qy	2271	TGTGACCTTGACTGTATCTCTCATCGACCGGTGGAGGTGAGTCTTACTGCTGTGC	2330
Db	2005	TGTGACCTTGACTGTATCTCTCATCGACCGGTGGAGGTGAGTCTTACTGCTGTGC	2064
Qy	2331	CCTCGGGCTCATCATTTTGTGTGTGAAAAAGAAAAAGAAACAAGGGCCCGC	2390
Db	2065	CCTCGGGCTCATCATTTTGTGTGTGAAAAAGAAAAAGAAACAAGGGCCCGC	2124
Qy	2391	TGTGGGTATCTAATGGCAATCAATTACTAGATGGCGAGCGCAAAAAAGTTCA	2450
Db	2125	TGTGGGTATCTAATGGCAATCAATTACTAGATGGCGCAAAAAAGTTCA	2183
Qy	2451	GAAAGGCGCAAGAGCAATGACTCCCATGTGTATGAGTCAATGAGACACCATGGTATA	2510
Db	2184	GAAAGGCGCAAGAGCAATGACTCCCATGTGTATGAGTCAATGAGACACCATGGTATA	2243
Qy	2511	TGGGCACTGTCTACAGAGATTCCAGCGGCTCTTCTCTGACGACAGAGTGAACCTTACG	2570
Db	2244	TGGGCACTGTCTACAGAGATTCCAGCGGCTCTTCTCTGACGACAGAGTGAACCTTACG	2303
Qy	2571	GCGGTTCAGGGGACACATGGGGGTCTGTCTCCCTCCCAACCATATGCTCCAGGGC	2630
Db	2304	GCGGTTCAGGGGACACATGGGGGTCTGTCTCCCTCCCAACCATATGCTCCAGGGC	2363
Qy	2631	CCCAACTGCAAGTGGGCACTGAGAGGCACTCTGCTCCCTCTGAGTCTGAGAG	2690
Db	2364	CCCAACTGCAAGTGGGCACTGAGAGGCACTCTGCTCCCTCTGAGTCTGAGAG	2423
Qy	2691	TGAACCGTACACCTTCTCCCATCCCAACATGGGGATGTAAAGCAGCAAGACAT	2750
Db	2424	TGAACCGTACACCTTCTCCCATCCCAACATGGGGATGTAAAGCAGCAAGACAT	2483
Qy	2751	TCCCTTACTGAACATCTCAGAGAGCCCATGAGGCCAGAGAAATTAATCTTCCAGAC	2810
Db	2484	TCCCTTACTGAACATCTCAGAGAGCCCATGAGGCCAGAGAAATTAATCTTCCAGAC	2543
Qy	2811	GCTTGTCTGAGTTTCTTAAGAGAGGGCACTGAGACACCGTCCGTTCTTAACAGAAA	2870
Db	2544	GCTTGTCTGAGTTTCTTAAGAGAGGGCACTGAGACACCGTCCGTTCTTAACAGAAA	2603
Qy	2871	TCTTAAAGAGAGAAATTAACAGAGAAAGACAGAGAGTTTCTCTGACACCGCAAC	2930
Db	2604	TCTTAAAGAGAGAAATTAACAGAGAAAGACAGAGAGTTTCTCTGACACCGCAAC	2663
Qy	2931	TTTCAATTTGCTCAGTGAATCTAATTTAAAGGCAAGACATTTGAAATGATGATTCATC	2990
Db	2664	TTTCAATTTGCTCAGTGAATCTAATTTAAAGGCAAGACATTTGAAATGATGATTCATC	2723
Qy	2991	TGATATACAGTCAATGACAGTCAATGCTCTCACTTAGAGCTGTGGGTTAGCAGCGTC	3050
Db	2724	TGATATACAGTCAATGACAGTCAATGCTCTCACTTAGAGCTGTGGGTTAGCAGCGTC	2783
Qy	3051	TAAATAGAGAGAGAGGCTGAGTCACTTAAGATAGGTTGACAGCAAGCCTGGAATTCAG	3110
Db	2784	TAAATAGAGAGAGAGGCTGAGTCACTTAAGATAGGTTGACAGCAAGCCTGGAATTCAG	2843
Qy	3111	AGTGTAAACAGAGGCTTGCCCTCTTTCAGAGCAACAGTTTCCAAATTCAGAGAGCTTACT	3170
Db	2844	AGTGTAAACAGAGGCTTGCCCTCTTTCAGAGCAACAGTTTCCAAATTCAGAGAGCTTACT	2903
Qy	3171	GAGGTCCCTTACTCTGACTGGGGTCCCAAGATGAAAAAGCAATGTGCTTTTATTAAT	3230
Db	2904	GAGGTCCCTTACTCTGACTGGGGTCCCAAGATGAAAAAGCAATGTGCTTTTATTAAT	2963
Qy	3231	ATTATATTGGTGTCTGTGTATTTAAAGAGATCAAAATGTATTAACCACTTACTTTTC	3290
Db	2964	ATTATATTGGTGTCTGTGTATTTAAAGAGATCAAAATGTATTAACCACTTACTTTTC	3023
Qy	3291	ACCTGACTTAAATTAATCACTTAACTTAACTGTTGAGATGAGCTGGGTGATCTTACTG	3350
Db	3024	ACCTGACTTAAATTAATCACTTAACTTAACTGTTGAGATGAGCTGGGTGATCTTACTG	3083
Qy	3351	ACCGTAAATTAACGTTGCTGCTGCTCCCAAGGTGGGATTAATTAACATCTGTCAA	3410
Db	3084	ACCGTAAATTAACGTTGCTGCTGCTCCCAAGGTGGGATTAATTAACATCTGTCAA	3143
Qy	3411	CCAGAAAAAGATGTGTGTGTTGAGCAGATTTGACATATCTGCTTGAATGAAGACTT	3470
Db	3144	CCAGAAAAAGATGTGTGTGTTGAGCAGATTTGACATATCTGCTTGAATGAAGACTT	3203
Qy	3471	CCTGATTTCTTAAAGTGGTGGTGTATCCCATTTGAGAAATTCATCTTGAATCCAT	3530
Db	3204	CCTGATTTCTTAAAGTGGTGGTGTATCCCATTTGAGAAATTCATCTTGAATCCAT	3263
Qy	3531	GTCTATAGTCTTAAAGATTAAGAAATTTCTCAAGTTTCCATGTGGGTTCTCTAGC	3590
Db	3264	GTCTATAGTCTTAAAGATTAAGAAATTTCTCAAGTTTCCATGTGGGTTCTCTAGC	3323
Qy	3591	TGCAGCAATCTTGAACATTTAAAGAAATTTAAGAAATTTCAATCTTCAAAATG	3650
Db	3324	TGCAGCAATCTTGAACATTTAAAGAAATTTAAGAAATTTCAATCTTCAAAATG	3383
Qy	3651	TTTAATTAATTAACAAAGTGGGCTGCTGATTTGCTTGTGCACTGCAACCAT	3710
Db	3384	TTTAATTAATTAACAAAGTGGGCTGCTGATTTGCTTGTGCACTGCAACCAT	3443
Qy	3711	TACTGTGATGCTTAAACCAACATTAAGTATAGTCTTGGGATTCAGAAATTCAGAAAT	3770
Db	3444	TACTGTGATGCTTAAACCAACATTAAGTATAGTCTTGGGATTCAGAAATTCAGAAAT	3503
Qy	3771	GGATGCTCCCTGAATGAATCAAGGTGTCAAGAGCTGTGCTCTTGAAGGCTTGA	3830
Db	3504	GGATGCTCCCTGAATGAATCAAGGTGTCAAGAGCTGTGCTCTTGAAGGCTTGA	3563
Qy	3831	GGAGAAAGCGGTTCTTCCATTTTCAAGCTTCTAAGGCTGGCTGCATTTCCAGGCTCA	3890
Db	3564	GGAGAAAGCGGTTCTTCCATTTTCAAGCTTCTAAGGCTGGCTGCATTTCCAGGCTCA	3623
Qy	3891	GTGGCTGATCAAGCTTTTCTCATGATGACCTGTGACATCGGCGCTCCCACTTCCCTC	3950
Db	3624	GTGGCTGATCAAGCTTTTCTCATGATGACCTGTGACATCGGCGCTCCCACTTCCCTC	3683
Qy	3951	TTTGAATTAACAAAGCCCAACAGAAAGATCCAGATTAATCTCTCAATTAAGATCTTCA	4010
Db	3684	TTTGAATTAACAAAGCCCAACAGAAAGATCCAGATTAATCTCTCAATTAAGATCTTCA	3743
Qy	4011	TCATCTTGAAGAGCTTTTTCATGACAGAAACAAATATAGCCACAGTGGGATTAAGACC	4070
Db	3744	TCATCTTGAAGAGCTTTTTCATGACAGAAACAAATATAGCCACAGTGGGATTAAGACC	3803
Qy	4071	AGGACATCTTTGGGGTGTGTTATCTGCTTCAACATCTTCTGCACTGACTCCCAAC	4130
Db	3804	AGGACATCTTTGGGGTGTGTTATCTGCTTCAACATCTTCTGCACTGACTCCCAAC	3863
Qy	4131	GGAGAGGCTAACAAATGATCTGGCGCACAGGGAATGTTTGTTAAGCTTGGAGCTTAAC	4190
Db	3864	GGAGAGGCTAACAAATGATCTGGCGCACAGGGAATGTTTGTTAAGCTTGGAGCTTAAC	3923

QY 4191 ACTTAAAAAACCAGATCAGAAATCTGGCCATCTGGGGCTCACTTTCACCTAGC 4250
DB 3924 ACTTAAAAAACCAGATCAGAAATCTGGCCATCTGGGGCTCACTTTCACCTAGC 3983
QY 4251 AACCACTGGGTGGAGTGGGCAACAGCTGCTTTAGAAAGGGGTTCACCTTCAACAG 4310
DB 3984 AACCACTGGGTGGAGTGGGCAACAGCTGCTTTAGAAAGGGGTTCACCTTCAACAG 4043
QY 4311 TCACCAAGCCCACTAGACCTTATCACTTCCCAATAGAGGCTAAGTGTGTTCTA 4370
DB 4044 TCACCAAGCCCACTAGACCTTATCACTTCCCAATAGAGGCTAAGTGTGTTCTA 4103
QY 4371 CTGATCAATGCCCTGAGGTGCTATTTATTTATGAAAAAAGAAAGACTGGATTAATC 4430
DB 4104 CTGATCAATGCCCTGAGGTGCTATTTATTTATGAAAAAAGAAAGACTGGATTAATC 4163
QY 4431 TCTAACTAGGTGATGAGCAATGAGCAATGTGTCTCACTTACCCTTTTCTTTT 4490
DB 4164 TCTAACTAGGTGATGAGCAATGAGCAATGTGTCTCACTTACCCTTTTCTTTT 4223
QY 4491 TTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 4550
DB 4224 TTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 4283
QY 4551 TGGAGTGCAGTGGCGCAATCTCGGCTCACTGCAACTCTGCTCTGGGCTCAAGCAAT 4610
DB 4284 TGGAGTGCAGTGGCGCAATCTCGGCTCACTGCAACTCTGCTCTGGGCTCAAGCAAT 4343
QY 4611 CTCCCACTCAGCTCCCAATATGCTGGGATCACTGGCAAAACCAACATGCCACCTAA 4670
DB 4344 CTCCCACTCAGCTCCCAATATGCTGGGATCACTGGCAAAACCAACATGCCACCTAA 4403
QY 4671 TTTTGTATTTTGTAGAGACAGGTTTCACTGTTGCCAGCTGCTCTCACTCTCT 4730
DB 4404 TTTTGTATTTTGTAGAGACAGGTTTCACTGTTGCCAGCTGCTCTCACTCTCT 4463
QY 4731 GGGCTCAAGCAATCTCTGCTCTGCTCTGCTCCAAAGTCTGGGATTAACAGATGTAGCA 4790
DB 4464 GGGCTCAAGCAATCTCTGCTCTGCTCTGCTCCAAAGTCTGGGATTAACAGATGTAGCA 4523
QY 4791 CCGCATCCAGCCCAACCTCTATTAACCAATTAACCTGCCCAAGTAACGTGGAATTTT 4850
DB 4524 CCGCATCCAGCCCAACCTCTATTAACCAATTAACCTGCCCAAGTAACGTGGAATTTT 4583
QY 4851 GCTTCTCAACCTGCTGCTGATCTGGAAGAGAGGATTAATGTTATAGCTTGCAGACA 4910
DB 4584 GCTTCTCAACCTGCTGCTGATCTGGAAGAGAGGATTAATGTTATAGCTTGCAGACA 4643
QY 4911 GTCCCAAGTTCAATATTTCTGGGCAAAACCTTCTTCAAAAAATTAATGTACTTATG 4970
DB 4644 GTCCCAAGTTCAATATTTCTGGGCAAAACCTTCTTCAAAAAATTAATGTACTTATG 4703
QY 4971 TATTCATGAATTCACCTTGGAAATGACCGCTCAACCTTGTCAATGGCAATAATGAA 5030
DB 4704 TATTCATGAATTCACCTTGGAAATGACCGCTCAACCTTGTCAATGGCAATAATGAA 4763
QY 5031 AGGAATTTTATAGTCTTAATAGGCTGATCTGCAAGACCTTCTTGAACACTTCCAGAG 5090
DB 4764 AGGAATTTTATAGTCTTAATAGGCTGATCTGCAAGACCTTCTTGAACACTTCCAGAG 4823
QY 5091 GATAGATATTTAATGATGATGCTTGGGCTTGGCTTATGGAACCTTCTTCTGAAAGTC 5150
DB 4824 GATAGATATTTAATGATGATGCTTGGGCTTGGCTTATGGAACCTTCTTCTGAAAGTC 4883
QY 5151 TGGTTCCTGCGCCAGTACCTTGGCTTGGAGCCAGATGCTGACCTTGCATTAAGGGC 5210
DB 4884 TGGTTCCTGCGCCAGTACCTTGGCTTGGAGCCAGATGCTGACCTTGCATTAAGGGC 4943
QY 5211 CAAGAAGAGGCTCGGCTTCTTCTCTCACTGAAGAGCCCTTATTTGAATTCAGTGTG 5270
DB 4944 CAAGAAGAGGCTCGGCTTCTTCTCTCACTGAAGAGCCCTTATTTGAATTCAGTGTG 5003

QY 5271 GAGCCCTAGCCCTCCATCTTGTGACATTCGCCAACTCCAGGCCCTTCCAAAGAGACTA 5330
DB 5004 GAGCCCTAGCCCTCCATCTTGTGACATTCGCCAACTCCAGGCCCTTCCAAAGAGACTA 5063
QY 5331 GGTGCCCTGCATTTCCACCAGAGTGGAGTTGGCTTCTCTAGAGCTGGCTACTTGTACCA 5390
DB 5064 GGTGCCCTGCATTTCCACCAGAGTGGAGTTGGCTTCTCTAGAGCTGGCTACTTGTACCA 5123
QY 5391 TCACCAATCTCTGAGCAAGTGTGCTGCTGAGACCAAGTATTTCTTGAAGCTGTGCTCA 5450
DB 5124 TCACCAATCTCTGAGCAAGTGTGCTGAGACCAAGTATTTCTTGAAGCTGTGCTCA 5183
QY 5451 TCAAAATCTCTGAGCAAGTGTGCTGCTGAGACCAAGTATTTCTTGAAGCTGTGCTCA 5510
DB 5184 TCAAAATCTCTGAGCAAGTGTGCTGCTGAGACCAAGTATTTCTTGAAGCTGTGCTCA 5243
QY 5511 GTGAAGGGGCCAGGCTGAGAACCTGGCTTTTCTTTTAAAGCCAGGCCCACTTAC 5570
DB 5244 GTGAAGGGGCCAGGCTGAGAACCTGGCTTTTCTTTTAAAGCCAGGCCCACTTAC 5303
QY 5571 ATAAAACATTTCAAGGATCACTGGAACAGTGAAGTGCATTTGTTGAAGCTTACATG 5630
DB 5304 ATAAAACATTTCAAGGATCACTGGAACAGTGAAGTGCATTTGTTGAAGCTTACATG 5363
QY 5631 CCAGCCCACTGCTCATTCACGTTGCTGCTGATGCTGAGAGAGGCCAGCCGATGACAG 5690
DB 5364 CCAGCCCACTGCTCATTCACGTTGCTGCTGATGCTGAGAGAGGCCAGCCGATGACAG 5423
QY 5691 ACTGGCTCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5750
DB 5424 ANTGGCTCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5483
QY 5751 GGAACAAGCAAGGCCCAACCGGACATGGCTTGTGTAAGGTTAGACACTGGTGTGTG 5810
DB 5484 GGAACAAGCAAGGCCCAACCGGACATGGCTTGTGTAAGGTTAGACACTGGTGTGTG 5543
QY 5811 GATTCGACAGTCTTCACTGGAATAATTTATTCATTCGACATCTTTTAAAGTGTGCA 5870
DB 5544 GATTCGACAGTCTTCACTGGAATAATTTATTCATTCGACATCTTTTAAAGTGTGCA 5603
QY 5871 TATTCATTTCTGCTGCTTAAATTAACAAATGATACCAAAAAACAGTATCAAGCTGTTA 5930
DB 5604 TATTCATTTCTGCTGCTTAAATTAACAAATGATACCAAAAAACAGTATCAAGCTGTTA 5663
QY 5931 AGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5990
DB 5664 AGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5723
QY 5991 TGAACAGCTCAGCATGGGCTCAGCAGATGCTGCTTAAATTTGTGATGATACGAAGCC 6050
DB 5724 TGAACAGCTCAGCATGGGCTCAGCAGATGCTGCTTAAATTTGTGATGATACGAAGCC 5783
QY 6051 AGGCTTTGGGATTAACAAGTTCTTCTCTTCAATTTAGTGCAGCTGCTGAGAGCAT 6110
DB 5784 AGGCTTTGGGATTAACAAGTTCTTCTCTTCAATTTAGTGCAGCTGCTGAGAGCAT 5843
QY 6111 GTTTTGTCCGGAATAAAAAATTAATAGTCTTGAAGTCTGCCAAAAAATTTT 6163
DB 5844 GTTTTGTCCGGAATAAAAAATTAATAGTCTTGAAGTCTGCCAAAAAATTTT 5896

RESULT 4
US-10-128-692A-161
; Sequence 161, Application US/10128692A
; Publication No. US20040009547A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.


```

Db      ||| 181 TTGAGATGCTCTGCGACGAGAAAGCAATTACAGTTCTATTAAGTGGGACCCCG 240
Qy      ||| 430 ACTCTGCTGGCAAAACCCCTGTTACATCGTCATTCTTAAAGACATATAACATGTTGTCC 489
Db      ||| 241 ACTCTGCTGGCAAAACCCCTGTTACATCGTCATTCTTAAAGACATATAACATGTTGTCC 300
Qy      ||| 490 ATCAAGTCTGAGAAAGAAATAGCTTTTACCTTTAGCTGCGAGAGTCTTGAGATCACTTT 549
Db      ||| 301 ATCAAGTCTGAGAAAGAAATAGCTTTTACCTTTAGCTGCGAGAGTCTTGAGATCACTTT 360
Qy      ||| 550 GTCATAGAGATCCAGAAAATAATGACTGTATGTGAGGCCCATGTCTCTTTGGGAGGTT 609
Db      ||| 361 GTCATAGAGATCCAGAAAATAATGACTGTATGTGAGGCCCATGTCTCTTTGGGAGGTT 420
Qy      ||| 610 CAGCTTCAGCCCTCGACATCGTTGTTGCTTACCCCTCAACAGAACTTTCACTGGGATGC 669
Db      ||| 421 CAGCTTCAGCCCTCGACATCGTTGTTGCTTACCCCTCAACAGAACTTTCACTGGGATGC 480
Qy      ||| 670 AAAGCTCATTAAGAGCATCGGTTTAAAGCTGAGTTTCCATCCCTGCGCTGAGGAGATC 729
Db      ||| 481 AAAGCTCATTAAGAGCATCGGTTTAAAGCTGAGTTTCCATCCCTGCGCTGAGGAGATC 540
Qy      ||| 730 GGTCCGGGTGAGAGCTGCCCAAGCGAGTCACTCACTCATCAAGCGGCGGAATCGATGCC 789
Db      ||| 541 GGTCCGGGTGAGAGCTGCCCAAGCGAGTCACTCACTCATCAAGCGGCGGAATCGATGCC 600
Qy      ||| 790 ACCGTGTGAGAGTGGGAACTTTCTGACAGCAATGGACATGTCTCCCGAGTCAAGATGCA 849
Db      ||| 601 ACCGTGTGAGAGTGGGAACTTTCTGACAGCAATGGACATGTCTCCCGAGTCAAGATGCA 660
Qy      ||| 850 GAAGAGTGAATAATGCTTTACCACTCCGATGTTTCAACCCAGAAATGTTCTCCGCTTC 909
Db      ||| 661 GAAGAGTGAATAATGCTTTACCACTCCGATGTTTCAACCCAGAAATGTTCTCCGCTTC 720
Qy      ||| 910 AGCATTTGCAAAACCGCTCATCTATAAAAGCTGTGTGATCATGAGTCTGTGTTGAGGTT 969
Db      ||| 721 AGCATTTGCAAAACCGCTCATCTATAAAAGCTGTGTGATCATGAGTCTGTGTTGAGGTT 780
Qy      ||| 970 GAAGGCTGAGCAACCTGATGTCTGCACTACCCAGAGGCTTCCCTGAGATGAGCTTC 1029
Db      ||| 781 GAAGGCTGAGCAACCTGATGTCTGCACTACCCAGAGGCTTCCCTGAGATGAGCTTC 840
Qy      ||| 1030 ATGACGTGAGCAATTTGCTCTCTGCAACCTGCGGGGCGAGGCTCTCTTCTCAACTTC 1089
Db      ||| 841 ATGACGTGAGCAATTTGCTCTCTGCAACCTGCGGGGCGAGGCTCTCTTCTCAACTTC 900
Qy      ||| 1090 AACCTCTCCAACTGTGAGAGAGAGAGCGGGTTGAATACTACAATCCCGGGCTCCACC 1149
Db      ||| 901 AACCTCTCCAACTGTGAGAGAGAGAGCGGGTTGAATACTACAATCCCGGGCTCCACC 960
Qy      ||| 1150 ACCCAACCCCGAGGTGTTCAAGCTGAGAGCAAGACAGCTCTGGGAACTATGCGGGGAACTTC 1209
Db      ||| 961 ACCCAACCCCGAGGTGTTCAAGCTGAGAGCAAGACAGCTCTGGGAACTATGCGGGGAACTTC 1020
Qy      ||| 1210 AACCTCTCTCTGCAAGGCTGTGACCAAGATGCCCAAGTCCAGGGATCTCTCCGGCTGCAG 1269
Db      ||| 1021 AACCTCTCTCTGCAAGGCTGTGACCAAGATGCCCAAGTCCAGGGATCTCTCCGGCTGCAG 1080
Qy      ||| 1270 TTCCAAGTTTGGTCCCAATCCCAAAATGAAGA 1306
Db      ||| 1081 TTCCAAGTTTGGTCCCAATCCCAAAATGAAGA 1117

```

RESULT 7
US-10-140-018-161

; Sequence 161, Application US/10140018
; Publication No. US20030138885A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc

```

; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P330R1C158
; CURRENT APPLICATION NUMBER: US/10/140,018
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 161
; LENGTH: 1376
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-10-140-018-161

Query Match      18.1%; Score 1117; DB 13; Length 1376;
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
Matches 1117; Conservative 0;

Qy      ||| 190 CCAAGGCGGAGAGGAGCTGCGCCCGGCTTGGGCGCTTGGGCGCCGCGCTCCCAACGTCG 249
Db      ||| 1 CCAAGGCGGAGAGGAGCTGCGCCCGGCTTGGGCGCTTGGGCGCCGCGCTCCCAACGTCG 60
Qy      ||| 250 TTTTCCCAACGAGGCGGAGGCGTCTCCGAGTCAATGCGCGGCTTAATCTGCGGGTCTCT 309
Db      ||| 61 TTTTCCCAACGAGGCGGAGGCGTCTCCGAGTCAATGCGCGGCTTAATCTGCGGGTCTCT 120
Qy      ||| 310 ATGCACTGTAGAGGGGTTCTGCTGCTGAGGAGGCGGCTGCGCGGCGGAGAGACT 369
Db      ||| 121 ATGCACTGTAGAGGGGTTCTGCTGCTGAGGAGGCGGCTGCGCGGCGGAGAGACT 180
Qy      ||| 370 TTGAGATTTGCTCTGCGCAAGAAAGCAATTAAGTTTCAATTAAGCTGGGAGCCCG 429
Db      ||| 181 TTGAGATTTGCTCTGCGCAAGAAAGCAATTAAGTTTCAATTAAGCTGGGAGCCCG 240
Qy      ||| 430 ACTCTGCTGGCAAAACCTGTTTACATCGTCAATTTCTAAAGACATATAACATGTTGTCC 489
Db      ||| 241 ACTCTGCTGGCAAAACCTGTTTACATCGTCAATTTCTAAAGACATATAACATGTTGTCC 300
Qy      ||| 490 ATCAAGTCTGAGAAAGAAATAGCTTTTACCTTTAGCTGCGAGAGTCTTGAGATCACTTT 549
Db      ||| 301 ATCAAGTCTGAGAAAGAAATAGCTTTTACCTTTAGCTGCGAGAGTCTTGAGATCACTTT 360
Qy      ||| 550 GTCATAGAGATCCAGAAAATAATGACTGTATGTGAGGCCCATGTCTCTTTGGGAGGTT 609
Db      ||| 361 GTCATAGAGATCCAGAAAATAATGACTGTATGTGAGGCCCATGTCTCTTTGGGAGGTT 420
Qy      ||| 610 CAGCTTCAGCCCTCGACATCGTTGTTGCTTACCCCTCAACAGAACTTTCACTGGGATGC 669
Db      ||| 421 CAGCTTCAGCCCTCGACATCGTTGTTGCTTACCCCTCAACAGAACTTTCACTGGGATGC 480
Qy      ||| 670 AAAGCTCATTAAGAGCATCGGTTTAAAGCTGAGTTTCCATCCCTGCGCTGAGGAGATC 729
Db      ||| 481 AAAGCTCATTAAGAGCATCGGTTTAAAGCTGAGTTTCCATCCCTGCGCTGAGGAGATC 540
Qy      ||| 730 GGTCCGGGTGAGAGCTGCCCAAGCGAGTCACTCACTCATCAAGCGGCGGAATCGATGCC 789
Db      ||| 541 GGTCCGGGTGAGAGCTGCCCAAGCGAGTCACTCACTCATCAAGCGGCGGAATCGATGCC 600
Qy      ||| 790 ACCGTGTGAGAGTGAAGACTTTCTGAGCAATGGCACTGTGTCCCGAGTCAAGATGCAA 849
Db      ||| 601 ACCGTGTGAGAGTGAAGACTTTCTGAGCAATGGCACTGTGTCCCGAGTCAAGATGCAA 660

```

QY 850 GAAGAGTGAATGCGCTTACACCTTCCAGTGTCCACCCGAAATGTCTCCGCTTC 909
DB 661 GAAGAGTGAATGCGCTTACACCTTCCAGTGTCCACCCGAAATGTCTCCGCTTC 720
QY 910 AGGATTGCAACCGCTCATCTATTAACCGCTGTGATCATCGAGTCTGTGTTGAGGGT 969
DB 721 AGCATTGCAACCGCTCATCTATTAACCGCTGTGATCATCGAGTCTGTGTTGAGGGT 780
QY 970 GAAGCTCAGCAACCCGTGATGTCTGCACTACCGAAGGCTTCCCTGAGATGAGCTC 1029
DB 781 GAAGGCTCAGCAACCCGTGATGTCTGCACTACCGAAGGCTTCCCTGAGATGAGCTC 840
QY 1030 ATGACGTGGAGTTTGTCTCTCTGCAACCTGCGGGCCAGGCTCTCTCTCAACTTC 1089
DB 841 ATGACGTGGAGTTTGTCTCTCTGCAACCTGCGGGCCAGGCTCTCTCTCAACTTC 900
QY 1090 AACCTCTCCAACTGTGAG 1149
DB 901 AACCTCTCCAACTGTGAG 960
QY 1150 ACCAACCCCGAGGTGTTCAAGCTGAGAGACAGACAGCTGGAAACATGGCGGGAATTC 1209
DB 961 ACCAACCCCGAGGTGTTCAAGCTGAGAGACAGACAGCTGGAAACATGGCGGGAATTC 1020
QY 1210 AACCTCTCTGCAAGGCTGTGAACCAAGATGCCAAAGTCCAGGATCTCTCGGCTGAG 1269
DB 1021 AACCTCTCTGCAAGGCTGTGAACCAAGATGCCAAAGTCCAGGATCTCTCGGCTGAG 1080
QY 1270 TTCCAGTTTGTGTCACATCCCAAAATGAAGA 1306
DB 1081 TTCCAGTTTGTGTCACATCCCAAAATGAAGA 1117

RESULT 8
US-10-140-021-161

Sequence 161, Application US/10140021
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P330R1C167
CURRENT APPLICATION NUMBER: US/10/140, 021
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 161
LENGTH: 1376
TYPE: DNA
ORGANISM: Homo Sapien
US-10-140-021-161

Query Match 18.1%; Score 1117; DB 13; Length 1376;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 CCAAGGCGAGAGCGAGCTGCGCCGGCTTGGGCGCTTGAGGCGCGCGCTCCACCGCTG 249
DB 1 CCAAGGCGAGAGCGAGCTGCGCCGGCTTGGGCGCTTGAGGCGCGCGCTCCACCGCTG 60
QY 250 TTTTCCCAACCGAGGCGAGGCGCTCCCGAGTCAATGGCGGCTGAACCTGGGGGCTCT 309
DB 61 TTTTCCCAACCGAGGCGAGGCGCTCCCGAGTCAATGGCGGCTGAACCTGGGGGCTCT 120
QY 310 ATCGACCTGTAGGGGTTCTGTGCTGAGGTGCGGCGCGCTGCGCGCGAGAGCT 369
DB 121 ATCGACCTGTAGGGGTTCTGTGCTGAGGTGCGGCGCGCTGCGCGCGAGAGCT 180
QY 370 TTTGAGATTGCTCTGCGACAGAGAAACAACTTAAGTTCTATTAAGCTGGGACCCG 429
DB 181 TTTGAGATTGCTCTGCGACAGAGAAACAACTTAAGTTCTATTAAGCTGGGACCCG 240
QY 430 ACTGTGCGGCAAAACCGCTGTACATGCTATTTCTAAAGCATATTAACATGTTGTC 489
DB 241 ACTGTGCGGCAAAACCGCTGTACATGCTATTTCTAAAGCATATTAACATGTTGTC 300
QY 490 ATCAAGCTGAGAAAGAAATAGTCTTTACCTTTAGCTGCGAGAGTCTGAGATCACTT 549
DB 301 ATCAAGCTGAGAAAGAAATAGTCTTTACCTTTAGCTGCGAGAGTCTGAGATCACTT 360
QY 550 GTCAATAGAGATCAAGAAATATTTAGCTGATGTCAAGGCCATGTCTTTGGGAGGTT 609
DB 361 GTCAATAGAGATCAAGAAATATTTAGCTGATGTCAAGGCCATGTCTTTGGGAGGTT 420
QY 610 CAGCTTCAGCCCTGACATCGTGTGCTGCTACCCCTCAACAGACTTTTCACTGGGATGTC 669
DB 421 CAGCTTCAGCCCTGACATCGTGTGCTGCTACCCCTCAACAGACTTTTCACTGGGATGTC 480
QY 670 AAAGCTCAATAGAGATGAGGCTTTAGAGCTGAGTTTCCATCCCTGCTGAGAGATC 729
DB 481 AAAGCTCAATAGAGATGAGGCTTTAGAGCTGAGTTTCCATCCCTGCTGAGAGATC 540
QY 730 GGTCCGGGTGAGAGCTGCCAGAGAGTCACTCACTTCATCAGCGGCGGAATGATGCC 789
DB 541 GGTCCGGGTGAGAGCTGCCAGAGAGTCACTCACTTCATCAGCGGCGGAATGATGCC 600
QY 790 ACCGTGTGAGATGAGAACTTCTGCAAGATGCACTGTCTCCCGATCAAGATGCA 849
DB 601 ACCGTGTGAGATGAGAACTTCTGCAAGATGCACTGTCTCCCGATCAAGATGCA 660
QY 850 GAAGAGTGAATAATGAGCTTACACCTCCATGTTCCACCCAGAAATGTCTCCGGCTTC 909
DB 661 GAAGAGTGAATAATGAGCTTACACCTCCATGTTCCACCCAGAAATGTCTCCGGCTTC 720
QY 910 AGCATTGCAACCGCTCATCTATAAAGCTGTGATCATGAGTCTGTGTTGAGGGT 969
DB 721 AGCATTGCAACCGCTCATCTATAAAGCTGTGATCATGAGTCTGTGTTGAGGGT 780
QY 970 GAAGGCTCAGCAACCCGTGATGTCTGCACTACCGAAGGCTTCCCTGAGATGAGCTC 1029
DB 781 GAAGGCTCAGCAACCCGTGATGTCTGCACTACCGAAGGCTTCCCTGAGATGAGCTC 840
QY 1030 ATGACGTGGAGTTTGTCTCTCTGCAACCTGCGGGCCAGGCTCTCTCTCAACTTC 1089
DB 841 ATGACGTGGAGTTTGTCTCTCTGCAACCTGCGGGCCAGGCTCTCTCTCAACTTC 900
QY 1090 AACCTCTCCAACTGTGAG 1149
DB 901 AACCTCTCCAACTGTGAG 960
QY 1150 ACCAACCCCGAGGTGTTCAAGCTGAGAGACAGACAGCTGGAAACATGGCGGGAATTC 1209
DB 961 ACCAACCCCGAGGTGTTCAAGCTGAGAGACAGACAGCTGGAAACATGGCGGGAATTC 1020
QY 1210 AACCTCTCTGCAAGGCTGTGAACCAAGATGCCAAAGTCCAGGATCTCTCGGCTGAG 1269
DB 1021 AACCTCTCTGCAAGGCTGTGAACCAAGATGCCAAAGTCCAGGATCTCTCGGCTGAG 1080
QY 1270 TTCCAGTTTGTGTCACATCCCAAAATGAAGA 1306

Db 1081 TTCCAAGTTTGTGTCACATCCACAAATGAAGCA 1117

RESULT 9
US-10-140-274-161

; Sequence 161, Application US/10140274
; Publication No. US20030143674A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C161
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 161
; LENGTH: 1376
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-274-161

Query Match 18.1%; Score 1117; DB 13; Length 1376;

Best Local Similarity 100.0%; Pred.No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 190 CCAGGCGGAGCGCAGCTGCGCGGCTTGGGCGCGCGGCTCCCGCAGCTG 249
1 CCAGGCGGAGCGCAGCTGCGCGGCTTGGGCGCGCGGCTCCCGCAGCTG 60
250 TTTTCCCAACCGAGCGCGGCTCCCGAGTCATGCGCGGCTGAACTGCGGGCTCT 309
61 TTTTCCCAACCGAGCGCGGCTCCCGAGTCATGCGCGGCTGAACTGCGGGCTCT 120
310 ATCGCACTGCTAGGGGTTTCTGCTGCTGGGTGGCGCGGCTCCCGCGGCGGAGAGCT 369
121 ATCGCACTGCTAGGGGTTTCTGCTGCTGGGTGGCGCGGCTCCCGCGGCGGAGAGCT 180
370 TTTGAGATTGCTCTGCGCAAGAGAAACATTAAGTTCTGATAAGCTGGGGAACCCG 429
181 TTTGAGATTGCTCTGCGCAAGAGAAACATTAAGTTCTGATAAGCTGGGGAACCCG 240
430 ACTCTGCTGGCAAAACCTGTTACATCGTCATTTTAAAGACATATAACATGTTGTC 489
241 ACTCTGCTGGCAAAACCTGTTACATCGTCATTTTAAAGACATATAACATGTTGTC 300
480 ATCAAGTCTGGAGAAAGATAGTCTTTTACCTTTAGCTGCAAGTCTGAGAAATCACTT 549
301 ATCAAGTCTGGAGAAAGATAGTCTTTTACCTTTAGCTGCAAGTCTGAGAAATCACTT 360
550 GTCATAGAGATCCAGAAATATATGACTATATGTCAGGCCCATGTCCTTTTGGGAGGTT 609
361 GTCATAGAGATCCAGAAATATATGACTATATGTCAGGCCCATGTCCTTTTGGGAGGTT 420
610 CAGCTTCAGCCCTCGACATCGTTGTTGCTTACCTCAACAGAACTTTCATCTGGAGTTC 669

Db 421 CAGCTTCAGCCCTCGACATCGTTGCTTACCTCAACAGAACTTTCATCTGGAGTTC 480
Qy 670 AAAGCTCATTAAGAGATCGGTTTAAAGCTGCAAGTTTTCATCCCTGCGCTGAGGAGATC 729
Db 481 AAAGCTCATTAAGAGATCGGTTTAAAGCTGCAAGTTTTCATCCCTGCGCTGAGGAGATC 540
Qy 730 GGTCCGGGTGAGAGCTGCGCAGACGAGTCACTCACTCCATCAGCGGCGGAATGATGTC 789
Db 541 GGTCCGGGTGAGAGCTGCGCAGACGAGTCACTCACTCCATCAGCGGCGGAATGATGTC 600
Qy 790 ACCGTGTCAGATATGGAACCTTCTGCAAGCAATGCACTGTGTCCCGATCAAGATGCA 849
Db 601 ACCGTGTCAGATATGGAACCTTCTGCAAGCAATGCACTGTGTCCCGATCAAGATGCA 660
Qy 850 GAAGAGTGAATAATGCGCTTACCTCCATGCTTCCAGCCAGAAATGTCTCCGAGTTC 909
Db 661 GAAGAGTGAATAATGCGCTTACCTCCATGCTTCCAGCCAGAAATGTCTCCGAGTTC 720
Qy 910 AGCATTTGCAAAACCGCTCATCTATAAAACGTCTGTCATCATGAGTCTGTGTTGAGGT 969
Db 721 AGCATTTGCAAAACCGCTCATCTATAAAACGTCTGTCATCATGAGTCTGTGTTGAGGT 780
Qy 970 GAAGCTCAGCAACCTTGAATGCTGCACTACCAAGAGCTTCCCTGAGATGAGTTC 1029
Db 781 GAAGCTCAGCAACCTTGAATGCTGCACTACCAAGAGCTTCCCTGAGATGAGTTC 840
Qy 1030 ATGAGTGGCAAGTTTGTGCTTCTGCAACCTGCGGCGGAGGCTCTCTCTCAACTTC 1089
Db 841 ATGAGTGGCAAGTTTGTGCTTCTGCAACCTGCGGCGGAGGCTCTCTCTCAACTTC 900
Qy 1090 AACCTCTCAACTGTGAGAGAGAGAGAGCGGTTGAATCTACATCCCGGCTCCACC 1149
Db 901 AACCTCTCAACTGTGAGAGAGAGAGAGCGGTTGAATCTACATCCCGGCTCCACC 960
Qy 1150 ACCAACCCTGAGGTGTTAAGCTGAGAGCAACAGCTCGGGAATGCGGGGAATTC 1209
Db 961 ACCAACCCTGAGGTGTTAAGCTGAGAGCAACAGCTCGGGAATGCGGGGAATTC 1020
Qy 1210 AACCTCTCTGCAAGGCTGAGACCAAGATGCCAAGTCCAGGAGTCTCCGGCTGAG 1269
Db 1021 AACCTCTCTGCAAGGCTGAGACCAAGATGCCAAGTCCAGGAGTCTCCGGCTGAG 1080
Qy 1270 TTCCAAGTTTGTGTCACATCCACAAATGAAGCA 1306
Db 1081 TTCCAAGTTTGTGTCACATCCACAAATGAAGCA 1117

RESULT 10
US-10-140-471-161

; Sequence 161, Application US/10140471
; Publication No. US20030138887A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C163
; CURRENT APPLICATION NUMBER: US/10/140,471

```

; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 161
; LENGTH: 1376
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-471-161

```

Query Match	18.1%	Score 1117	DB 13	Length 1376
Best Local Similarity	100.0%	Pred. No.	0	
Matches 117	Conservative	0	Mismatches	0
			Indels	0
			Gaps	0

QY	190	CCAGGCGCGACCGCAGCTGCGCCGGGCTTTGGCGCTGGGAGCGCGCGCTCCCAACGATG	249
Db	1	CCAGGCGCGAAGCGAGCTGCGCCGGGCTTTGGCGCTGGGAGCGCGCGCTCCCAACGATG	60
QY	250	TTTTTCCCACCGAGGCGGAGCGTCCGAGATCATGGCCGGCTTAACTGCGGGTCTCT	309
Db	61	TTTTTCCCACCGAGGCGGAGCGTCCCGGAGCATGGCCGGCTTAACTGCGGGTCTCT	120
QY	310	ATGCCACTGCTAGGGGTTCTGTGTGGGTGGGGCGCTGCGCGGGGGCAAGT	369
Db	121	ATGCCACTGCTAGGGGTTCTGTGTGGGTGGGGCGCTGCGCGGGGGCAAGT	180
QY	370	TTTGAATTTGCTCTGCGCAGAGAAAGCAATTAAGTTCTATAAAGCTGGGGA	429
Db	181	TTTGAATTTGCTCTGCGCAGAGAAAGCAATTAAGTTCTATAAAGCTGGGGA	240
QY	430	ACTCTGCTGGCAAAACCTGTTACATGTCATTTCTAAAGCATATACATGTTGTCC	489
Db	241	ACTCTGCTGGCAAAACCTGTTACATGTCATTTCTAAAGCATATACATGTTGTCC	300
QY	490	ATCAAGTCGAGAAAGATAGTCTTTACCTTTAGCTGCAAGTCTGAGAAATCACTT	549
Db	301	ATCAAGTCGAGAAAGATAGTCTTTACCTTTAGCTGCAAGTCTGAGAAATCACTT	360
QY	550	GTCATAGCATCCAGAAAAATTGACCTGTATGTCAAGGCCCATGTCTTTGGGAGTT	609
Db	361	GTCATAGCATCCAGAAAAATTGACCTGTATGTCAAGGCCCATGTCTTTGGGAGTT	420
QY	610	CAGCTTCAGCCCTGCAGATCGTGTGGCTACCCCAACAGAACTTTACCTGGGATGTC	669
Db	421	CAGCTTCAGCCCTGCAGATCGTGTGGCTACCCCAACAGAACTTTACCTGGGATGTC	480
QY	670	AAAGCTCAAGAAGCATGCTTTAGCTGCACTTTCCATCCCTCGCCCTGAGGAGATC	729
Db	481	AAAGCTCAAGAAGCATGCTTTAGCTGCACTTTCCATCCCTCGCCCTGAGGAGATC	540
QY	730	GGTCCGGGTGAGAGCTGCGCAGAGGATCATCTCATAGGGCGGAAATGCATGTC	789
Db	541	GGTCCGGGTGAGAGCTGCGCAGAGGATCATCTCATAGGGCGGAAATGCATGTC	600
QY	790	ACCGTGTGACAGATCGGAACCTTCTGCAGCATGCACTGTGTCCCGATCAAGATGCA	849
Db	601	ACCGTGTGACAGATCGGAACCTTCTGCAGCATGCACTGTGTCCCGATCAAGATGCA	660
QY	850	GAAAGAGTGAATAATGCGCTTACACTCTCCATGCTTCCACCCAGAAATGTCTCCGGCTTC	909
Db	661	GAAAGAGTGAATAATGCGCTTACACTCTCCATGCTTCCACCCAGAAATGTCTCCGGCTTC	720
QY	910	AGCATTTGCAAAACCGCTCATCTATAAAGCGTGTGTCATCATGAGTCTGTGTTGAGGT	969
Db	721	AGCATTTGCAAAACCGCTCATCTATAAAGCGTGTGTCATCATGAGTCTGTGTTGAGGT	780
QY	970	GAAAGCTCAGCAACCTGATGTCTGCAACATACCAAGAAAGCTTCCCTGAGATGAGCTC	1022
Db	781	GAAAGCTCAGCAACCTGATGTCTGCAACATACCAAGAAAGCTTCCCTGAGATGAGCTC	840
QY	1030	ATGACGTGAGCATTTGTGTTCTGTCAGCAACCTGGGGGCGAGGTCCTTCTCTCAACTTC	1088
Db	841	ATGACGTGAGCATTTGTGTTCTGTCAGCAACCTGGGGGCGAGGTCCTTCTCTCAACTTC	900

Qy	1090	AACTCTCCAACTGTGAGAGAAAGAGACCGGGTTGAATACATCAATCCCGGGCTCCACC	1149
Db	901	AACTCTCCAACTGTGAGAGAAAGAGACCGGGTTGAATACATCAATCCCGGGCTCCACC	960
Qy	1150	ACCAACCCCGAGGTTTCAAGCTTGAGAGCAAGCAGCTTGGAACATGCGCGGGAACCTTC	1209
Db	961	ACCAACCCCGAGGTTTCAAGCTTGAGAGCAAGCAGCTTGGAACATGCGCGGGAACCTTC	1020
Qy	1210	AACTCTCTCTGCAAGGCTGTGACCAAGATGCCCCAAGTCCAGGGAATCCTTCGGCTCAG	1269
Db	1021	AACTCTCTCTGCAAGGCTGTGACCAAGATGCCCCAAGTCCAGGGAATCCTTCGGCTCAG	1080
Qy	1270	TTCCAGTTTGGTCCAAATCCCAAAATGAAAGCA	1306
Db	1081	TTCCAGTTTGGTCCAAATCCCAAAATGAAAGCA	1117

```

RESULT 11
US-10-140-807-161
; Sequence 161, Application US/10140807
; Publication No. US20030134354A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C174
; CURRENT APPLICATION NUMBER: US/10/140,807
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 161
; LENGTH: 1376
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-807-161

```

Query Match	18.1%	Score 1117;	DB 13;	Length 1376;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 117; Conservative	0;	Mismatches	0;	Indels 0;
				Gaps 0;

[illegible]

```
Db 241 ACTCTGCTGGCAAAACCCCTGTATACCTGCTATTTCTAAAGACATATAACATGTTGTC 300
Qy 490 ATCAAGTCTGGAGAAAGATAGTCTTTACCTTAGCTGCAGAGTCCGAGAAATCACTTT 549
Db 301 ATCAAGTCTGGAGAAAGATAGTCTTTACCTTAGCTGCAGAGTCCGAGAAATCACTTT 360
Qy 550 GTCATAGAGATCCAGAAAAATATTGACTATATGTCAGGCCCATGTCCTTTGGGAGGTT 609
Db 361 GTCATAGAGATCCAGAAAAATATTGACTATATGTCAGGCCCATGTCCTTTGGGAGGTT 420
Qy 610 CAGCTTCAGCCCTCGACATGTTGTTGCTTACCTTCACAGAACTTTCACTCGGAGTGC 669
Db 421 CAGCTTCAGCCCTCGACATGTTGTTGCTTACCTTCACAGAACTTTCACTCGGAGTGC 480
Qy 670 AAAGCTCATAGAGATCGGTTTAAAGCTGCACTGCTTTCCATCCCTGCTGAGCAAGTTC 729
Db 481 AAAGCTCATAGAGATCGGTTTAAAGCTGCACTGCTTTCCATCCCTGCTGAGCAAGTTC 540
Qy 730 GGTCCGGGTGAGAGTCCAGAGAGTCACTCACTCCATCAGCGGCGGAATCGATGCC 789
Db 541 GGTCCGGGTGAGAGTCCAGAGAGTCACTCACTCCATCAGCGGCGGAATCGATGCC 600
Qy 790 ACCGTGCTCAGAGTGGAACTTTCTGCAAGCAATGGCACTGTGTCGAGTCAAGATGCA 849
Db 601 ACCGTGCTCAGAGTGGAACTTTCTGCAAGCAATGGCACTGTGTCGAGTCAAGATGCA 660
Qy 850 GAAGAGTGAAGAAATGAGCTTTACACTCCCATGTTCCACCCAGAAATGTCTCCGGCTTC 909
Db 661 GAAGAGTGAAGAAATGAGCTTTACACTCCCATGTTCCACCCAGAAATGTCTCCGGCTTC 720
Qy 910 AGCATTTGCAAAACCGCTCATCTATAAAGTCTGTGATCATGAGTGTGTTAGAGGT 969
Db 721 AGCATTTGCAAAACCGCTCATCTATAAAGTCTGTGATCATGAGTGTGTTAGAGGT 780
Qy 970 GAAGGCTCAGCAACCTGATGTCTGCAACTACCCAGAGGCTTCCCTGAGGATGAGTTC 1029
Db 781 GAAGGCTCAGCAACCTGATGTCTGCAACTACCCAGAGGCTTCCCTGAGGATGAGTTC 840
Qy 1030 ATGAGCTGAGCAAGTGTGCTTCTGCAACTGCTGCGGCGGAGGCTCTCTTCTCAACTTC 1089
Db 841 ATGAGCTGAGCAAGTGTGCTTCTGCAACTGCTGCGGCGGAGGCTCTCTTCTCAACTTC 900
Qy 1090 AACCTCTCCAACTGTGAGAGAGAGAGAGGAGGTTGAATCTACATCCCGGCTCCACC 1149
Db 901 AACCTCTCCAACTGTGAGAGAGAGAGAGGAGGTTGAATCTACATCCCGGCTCCACC 960
Qy 1150 ACCAACCCTGAGGTTTCAAGCTGAGAGAGAGAGAGGCTGAGAACATGAGCGGAGACTTC 1209
Db 961 ACCAACCCTGAGGTTTCAAGCTGAGAGAGAGAGAGGCTGAGAACATGAGCGGAGACTTC 1020
Qy 1210 AACCTCTCTCTGCAAGAGCTGTGACCAAGATGCTCCAAAGTCCAGGAGTCTCCGGCTGCAG 1269
Db 1021 AACCTCTCTCTGCAAGAGCTGTGACCAAGATGCTCCAAAGTCCAGGAGTCTCCGGCTGCAG 1080
Qy 1270 TTCCAAGTTTGGTCCAAATCCCAAAATGGAAGCA 1306
Db 1081 TTCCAAGTTTGGTCCAAATCCCAAAATGGAAGCA 1117
```

RESULT 12
US-10-140-922-161

```
Sequence 161, Application US/10140922
Publicatation No. US2003013889A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Geritsen, Mary E.
APPLICANT: Goddard, Audrey
```

```
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C179
CURRENT APPLICATION NUMBER: US/10/140,922
CURRENT FILING DATE: 2002-05-07
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 161
LENGTH: 1376
TYPE: DNA
ORGANISM: Homo Sapien
US-10-140-922-161

Query Match 18.1%; Score 1117; DB 13; Length 1376;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 190 CCAGGGCCGAGAGCGAGCTGCGGCTTTGGGCGCTTGGGCGCGGCTTCCCAACCTGCG 249
Db 1 CCAGGGCCGAGAGCGAGCTGCGGCTTTGGGCGCTTGGGCGCGGCTTCCCAACCTGCG 60
Qy 250 TTTTCCCAACCGAGGCGAGGCGTCCGAGATCATGCGCGGCTGAACTGCGGAGTCTCT 309
Db 61 TTTTCCCAACCGAGGCGAGGCGTCCGAGATCATGCGCGGCTGAACTGCGGAGTCTCT 120
Qy 310 ATGCACTGTAGAGGTTCTGCTGCTGAGGCGGCGGCTTCCGCGCGGAGAGAACT 369
Db 121 ATGCACTGTAGAGGTTCTGCTGCTGAGGCGGCGGCTTCCGCGCGGAGAGAACT 180
Qy 370 TTTGAGATTTGCTGCGGCGAGAGAAAGCAATTAAGTTTCAATAAAGTGGGAGCCCCG 429
Db 191 TTTGAGATTTGCTGCGGCGAGAGAAAGCAATTAAGTTTCAATAAAGTGGGAGCCCCG 240
Qy 430 ACTGTGCTGGCAAAACCTGTTTACATGCTCAATTTCTAAAGACATATAACATGTTGTC 489
Db 241 ACTGTGCTGGCAAAACCTGTTTACATGCTCAATTTCTAAAGACATATAACATGTTGTC 300
Qy 490 ATCAAGTCTGAGAAAGATAGTCTTTACCTTAGCTGCAGAGTCCGAGAAATCACTTT 549
Db 301 ATCAAGTCTGAGAAAGATAGTCTTTACCTTAGCTGCAGAGTCCGAGAAATCACTTT 360
Qy 550 GTCATAGAGATCCAGAAAAATATTGACTATATGTCAGGCCCATGTCCTTTGGGAGGTT 609
Db 361 GTCATAGAGATCCAGAAAAATATTGACTATATGTCAGGCCCATGTCCTTTGGGAGGTT 420
Qy 610 CAGCTTCAGCCCTCGACATGTTGTTGCTTACCTTCACAGAACTTTCACTCGGAGTGC 669
Db 421 CAGCTTCAGCCCTCGACATGTTGTTGCTTACCTTCACAGAACTTTCACTCGGAGTGC 480
Qy 670 AAAGCTCATAGAGATCGGTTTAAAGCTGCACTGCTTTCCATCCCTGCTGAGCAAGTTC 729
Db 481 AAAGCTCATAGAGATCGGTTTAAAGCTGCACTGCTTTCCATCCCTGCTGAGCAAGTTC 540
Qy 730 GGTCCGGGTGAGAGTCCAGAGAGTCACTCACTCCATCAGCGGCGGAATCGATGCC 789
Db 541 GGTCCGGGTGAGAGTCCAGAGAGTCACTCACTCCATCAGCGGCGGAATCGATGCC 600
Qy 790 ACCGTGCTCAGAGTGGAACTTTCTGCAAGCAATGGCACTGTGTCGAGTCAAGATGCA 849
Db 601 ACCGTGCTCAGAGTGGAACTTTCTGCAAGCAATGGCACTGTGTCGAGTCAAGATGCA 660
Qy 850 GAAGAGTGAAGAAATGAGCTTTACACTCCCATGTTCCACCCAGAAATGTCTCCGGCTTC 909
Db 661 GAAGAGTGAAGAAATGAGCTTTACACTCCCATGTTCCACCCAGAAATGTCTCCGGCTTC 720
```


LENGTH: 1376
TYPE: DNA
ORGANISM: Homo Sapien
US-10-141-698-161

Query Match 18.1% Score 1117; DB 13; Length 1376;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

190 CCAGGCGGAGCGCAGCTGCGCCGGGCTTGGGCGCTGGGGCGCGGCTCCACCACCGTGC 249
1 CCAGGCGGAGCGCAGCTGCGCCGGGCTTGGGCGCTGGGGCGCGGCTCCACCACCGTGC 60
250 TTTTCCCAACCGAGCGCGGCTCCCGAGTCAATGCGCGCTGAATGCGGGGCTCTT 309
61 TTTTCCCAACCGAGCGCGGCTCCCGAGTCAATGCGCGCTGAATGCGGGGCTCTT 120
310 ATCGCATGCTAGGGGCTCTGCTGCTGGGCTGGGCGCGGCTGGGCGCGGCGAGAGCT 369
121 ATCGCATGCTAGGGGCTCTGCTGCTGGGCTGGGCGCGGCTGGGCGCGGCGAGAGCT 180
370 TTTGAGATTGCTCTGCGCAGAGAAAGCAATTACAGTTCTGATAAAGCTGGGGAACCCG 429
181 TTTGAGATTGCTCTGCGCAGAGAAAGCAATTACAGTTCTGATAAAGCTGGGGAACCCG 240
430 ACTCTGCTGGCAAAACCTGTTACATGTCATTTCTAAAGACATTAACATGTTGTC 489
241 ACTCTGCTGGCAAAACCTGTTACATGTCATTTCTAAAGACATTAACATGTTGTC 300
490 ATCAAGTCTGGAAGAAATAGCTTACCTTTAGCTGCGAGAGCTGAGAAATCACTT 549
301 ATCAAGTCTGGAAGAAATAGCTTACCTTTAGCTGCGAGAGCTGAGAAATCACTT 360
550 GTCATAGAGATCCAGAAATAATGACTGATGTCAGGCCCATGTCCTTTTGGGAGATT 609
361 GTCATAGAGATCCAGAAATAATGACTGATGTCAGGCCCATGTCCTTTTGGGAGATT 420
610 CAGCTTCAGCCCTCGACATGCTGTTGCTCCTACCTCAACAGAACTTTCATCTGGAGTGC 669
421 CAGCTTCAGCCCTCGACATGCTGTTGCTCCTACCTCAACAGAACTTTCATCTGGAGTGC 480
670 AAAGCTCATTAAGAGCATGGGTTTAAAGCTGAGGCTTTTCAATCCCTGAGGCGAGATC 729
481 AAAGCTCATTAAGAGCATGGGTTTAAAGCTGAGGCTTTTCAATCCCTGAGGCGAGATC 540
730 GGTCCGGGTGAGAGCTGCCAGAGCTCACTCACTCATCAGCGGCGGAAATCGATGCC 789
541 GGTCCGGGTGAGAGCTGCCAGAGCTCACTCACTCATCAGCGGCGGAAATCGATGCC 600
790 ACCGTGCTCAGATCGGAACCTTCTGAGCAATGCGACTGTGCTCCGATCAAGATGCA 849
601 ACCGTGCTCAGATCGGAACCTTCTGAGCAATGCGACTGTGCTCCGATCAAGATGCA 660
850 GAAAGGATGAAATGGGCTTACACCTCCAGATGTTCAACCCAGAAATGTCCTCCGCTTC 909
661 GAAAGGATGAAATGGGCTTACACCTCCAGATGTTCAACCCAGAAATGTCCTCCGCTTC 720
910 AGCATTTGCAAAACCGCTCATCTAATAAAGCTGTGTCATCATGAGTCTGTGTTGAGGAT 969
721 AGCATTTGCAAAACCGCTCATCTAATAAAGCTGTGTCATCATGAGTCTGTGTTGAGGAT 780
970 GAAAGCTCAGCAACCTGATGTCGCAACTAACCCAGAAAGCTTCCCTGAGATGAGCTC 1029
781 GAAAGCTCAGCAACCTGATGTCGCAACTAACCCAGAAAGCTTCCCTGAGATGAGCTC 840
1030 ATGACCTGCGAGTTTGTCTTCTGCAACCTGCGGGCGAGGCTCTCTTCTCAACTTC 1089
841 ATGACCTGCGAGTTTGTCTTCTGCAACCTGCGGGCGAGGCTCTCTTCTCAACTTC 900
1090 AACCTCTCAACTGTGAGAGAGAGAGAGGAGGTTGAATACTACATCCCGGGCTCCACC 1149
901 AACCTCTCAACTGTGAGAGAGAGAGAGGAGGTTGAATACTACATCCCGGGCTCCACC 960

QY 1150 ACCAACCAGGCTGTTCAAGCTGAGAGCAAGAGCCTGGGAAACATGCGGGGAACTTC 1209
DB 961 ACCAACCAGGCTGTTCAAGCTGAGAGCAAGAGCCTGGGAAACATGCGGGGAACTTC 1020
QY 1210 AACCTCTCTGCAAGGCTGTGACCAAGATGCCAAAGTCCAGGATCCTCCGGCTGCAG 1269
DB 1021 AACCTCTCTGCAAGGCTGTGACCAAGATGCCAAAGTCCAGGATCCTCCGGCTGCAG 1080
QY 1270 TTCCAAGTTTGTGTCACCAATCCACCAAAATGAAAGCA 1306
DB 1081 TTCCAAGTTTGTGTCACCAATCCACCAAAATGAAAGCA 1117

Search completed: February 20, 2004, 19:44:59
Job time: 1986 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: February 20, 2004, 13:49:26 ; Search time 87 Seconds

(without alignments)
1525.235 Million cell updates/sec

Title: US-09-899-569a-4

Perfect score: 836
Sequence: 1 MAGLNCVSLALGVLLGA.....SSKQDIPLNTQBPMPAP 836

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size: 0
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

A_GeneSeq_19Jun03:*

1: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	836	100.0	836	23	ABP69553
2	836	100.0	836	23	AA49641
3	650	77.8	749	23	AA49640
4	413	49.4	443	21	AAV91592
5	382	45.7	649	22	AA85558
6	341	40.8	343	22	AAU1252
7	341	40.8	343	24	ABU6650
8	341	40.8	343	24	ABU6926
9	341	40.8	343	24	ABU5931

10	211	25.2	709	21	AAV91456
11	159	19.0	159	21	AAV91593
12	143	17.1	143	22	ABG00296
13	121	14.5	121	22	AA23969
14	74	8.9	116	22	AA23969
15	39	4.7	167	22	AAU69483
16	9	1.1	19	21	AAV69781
17	9	1.1	75	24	ABP56973
18	9	1.1	215	24	ABP71325
19	9	1.1	215	24	ABP56994
20	9	1.1	293	24	ABP57488
21	9	1.1	2359	24	ABP56959
22	9	1.1	2360	24	ABP56972
23	9	1.1	2661	24	ABP56958
24	8	1.0	62	23	ABP35396
25	8	1.0	70	24	ABP56985
26	8	1.0	71	20	AAV30830
27	8	1.0	75	23	ABP10049
28	8	1.0	78	22	ABG56028
29	8	1.0	78	22	ABG40578
30	8	1.0	78	22	ABG24869
31	8	1.0	78	22	AA651436
32	8	1.0	78	22	AAV74226
33	8	1.0	78	22	AAV34338
34	8	1.0	78	22	ABG44180
35	8	1.0	132	22	AAO00490
36	8	1.0	133	22	ABG22899
37	8	1.0	147	22	AAE03940
38	8	1.0	156	22	AAU18521
39	8	1.0	179	22	AA831192
40	8	1.0	187	22	AAU98984
41	8	1.0	193	22	AAU48107
42	8	1.0	213	21	AA818289
43	8	1.0	221	20	AA88495
44	8	1.0	221	21	AA825793
45	8	1.0	221	21	AAV32203

ALIGNMENTS

RESULT 1	ABP69553	ABP69553 standard; Protein; 836 AA.
ID	ABP69553	
AC	ABP69553	
DT	20-JAN-2003	(first entry)
XX		
DE	Human polypeptide SEQ ID NO 1600.	
XX		
KW	Human; genome mapping; gene therapy; food supplement; virus; fungus; cell-proliferative disorder; neurodegenerative disease; bacterial; Parkinson's disease; Alzheimer's disease; autoimmune disease; multiple sclerosis; diabetes; genetic disorder; wound; burn; infection; arthritis; cytostatic; immunomodulator; nootropic; neuroprotective; antiparkinsonian; antidiabetic; immunosuppressive; dermatological; haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide; antiarthritic.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200270539-A2.	
XX		
PD	12-SEP-2002.	
XX		
PF	05-MAR-2002; 2002WO-US05095.	
XX		
PR	05-MAR-2001; 2001US-0799451.	
XX		
PA	(HYSE-) HYSEQ INC.	
XX		
PI	Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;	

Human secreted pro
Human secreted pro
Novel human diagn
Human EST encoded
Human immune/haema
Human purified sec
Human tumour metas
E. maxima immuno
E. maxima 250 kDa
Mycobacterium tube
E. maxima 250 kDa
E. maxima immuno
Eimeria maxima 250
Human ORF4369 prot
Sarcocystis muris
Human secreted pro
Human ORF protein
Human liver peptid
Peptide #8084 enco
Protein #6868 enco
Human brain expres
Human bone marrow
Peptide #8375 enco
Human peptide enco
Human stromal cell
E. coli growth and
Propionibacterium
Plasmodium falci
Human stomach can
Human secreted pro
Human receptor mol

PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;
DR N-PSDB; AB211770.
XX MPI: 2002-759812/82.
XX
XX New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for treating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or
PT platelet or coagulation disorders -
XX
XX
PS Claim 9; SEQ ID NO 1600; 1012bp + Sequence Listing; English.
XX
XX The invention relates to an isolated polynucleotide (1) comprising a
CC nucleotide sequence selected from any of 948 sequences
CC (AB21119-AB212066) or their mature protein coding portion, active domain
CC coding protein or complementary sequences. The polynucleotides are useful
CC for identifying expressed genes or for physical mapping of human genome.
CC The encoded polypeptides (ABP68902-ABP69849) are useful as molecular
CC weight markers, as a food supplement, for generating antibodies, in
CC medical imaging, screening and diagnostic assays and for treating
CC cell-proliferative disorders (cancer), neurodegenerative diseases
CC (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple
CC sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid
CC disorders, platelet or coagulation disorders, wound, burns, incision,
CC ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,
CC parasitic), arthritis, etc.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 836 AA;

Query Match 100.0%; Score 836; DB 23; Length 836;
Best Local Similarity 100.0%; Pred No. 0;
Matches 836; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCVSIALGLVLLGAAALPRGAFAEIALPRESNITVLIKIGTPTLAKPCYIV 60
DB 1 MAGLNCVSIALGLVLLGAAALPRGAFAEIALPRESNITVLIKIGTPTLAKPCYIV 60
QY 61 SKRHITMLSIKSGERIVFTFSCSPENHVFIEIQKIDCMGCPGEGVQLQPSISLPT 120
DB 61 SKRHITMLSIKSGERIVFTFSCSPENHVFIEIQKIDCMGCPGEGVQLQPSISLPT 120
QY 121 LNRFTIWDVKAHSGIGLQPSIPRLQIGPESCPDGVTHSISGIDATVARIIGFCSN 180
DB 121 LNRFTIWDVKAHSGIGLQPSIPRLQIGPESCPDGVTHSISGIDATVARIIGFCSN 180
QY 181 GTVSRIKMGEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESVFEBSGATLMSANY 240
DB 181 GTVSRIKMGEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESVFEBSGATLMSANY 240
QY 241 PGCFPEDELMTWQFVPAHLRASVFLNENLSCGERKEVEYYTIGSTTNPVFLBPK 300
DB 241 PGCFPEDELMTWQFVPAHLRASVFLNENLSCGERKEVEYYTIGSTTNPVFLBPK 300
QY 241 PGCFPEDELMTWQFVPAHLRASVFLNENLSCGERKEVEYYTIGSTTNPVFLBPK 300
DB 241 PGCFPEDELMTWQFVPAHLRASVFLNENLSCGERKEVEYYTIGSTTNPVFLBPK 300
QY 301 QPENMAGNFPLSIQGCDDAQSPGILRLQVAVQHPQNSNKIYVVDLSNERAMSLTIE 360
DB 301 QPENMAGNFPLSIQGCDDAQSPGILRLQVAVQHPQNSNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVKSRKFPVPCFVLESRTCSNLTITSGSKHISFLCDLTLMMNVEKTIISCTH 420
DB 361 PRPVKSRKFPVPCFVLESRTCSNLTITSGSKHISFLCDLTLMMNVEKTIISCTH 420
QY 421 RYCQKRSYSIQVPSDILHPELHDFPSKVLVPRKRLSLVLPVQKLQOHTHEKPCNTSF 480
DB 421 RYCQKRSYSIQVPSDILHPELHDFPSKVLVPRKRLSLVLPVQKLQOHTHEKPCNTSF 480
QY 481 SYLVAAAIHQDLYFSGFCFGSGSIKQIQVKNISVLTFTAPBFOQBASQGLTVSFITF 540
DB 481 SYLVAAAIHQDLYFSGFCFGSGSIKQIQVKNISVLTFTAPBFOQBASQGLTVSFITF 540

QY 541 FKESGVFTVPTDTSKVLKIPNMDRGLPSLTUSWNISVPPDQVACLTFFKERSGVQC 600
DB 541 FKESGVFTVPTDTSKVLKIPNMDRGLPSLTUSWNISVPPDQVACLTFFKERSGVQC 600
QY 601 TGRAFMIIOEORTKAEELFSLDEVDLPKPSFHHHSFWNINISCSPTSGKQDLTSLVLT 660
DB 601 TGRAFMIIOEORTKAEELFSLDEVDLPKPSFHHHSFWNINISCSPTSGKQDLTSLVLT 660
QY 661 PRTVDLTVILAAVGGVLLSALGLIICVAKKKKKTKNKGPAVGIYNGNINTEMPOPK 720
DB 661 PRTVDLTVILAAVGGVLLSALGLIICVAKKKKKTKNKGPAVGIYNGNINTEMPOPK 720
QY 721 KFOGRKNDSHVAVIEDTWYGHLLQDSSGSPLOPEVDIYRPPQGTMGVCPSPPTIC 780
DB 721 KFOGRKNDSHVAVIEDTWYGHLLQDSSGSPLOPEVDIYRPPQGTMGVCPSPPTIC 780
QY 781 SRAPTAKLATEEPPRSPSESEPYTSHPNNGVSSKDNDIPLNTOEPMEPAE 836
DB 781 SRAPTAKLATEEPPRSPSESEPYTSHPNNGVSSKDNDIPLNTOEPMEPAE 836

RESULT 2
AAM49641
ID AAM49641 standard; Protein; 836 AA.
XX
AC AAM49641;
XX
DT 17-MAY-2002 (first entry)
DE Human tumour-associated antigen B345 protein SEQ ID NO 4.
XX
XX Tumour-associated antigen; human; B345; cytostatic; cell communication;
KW cell interaction; signal transduction; metastasis; cancer; colon;
KW immunotherapy; carcinoma; lung; diagnosis.
XX
OS Homo sapiens.
XX
PN WO200204508-A1.
XX
PD 17-JAN-2002.
XX
PF 05-JUL-2001; 2001WO-EP07705.
XX
PR 07-JUL-2000; 2000DE-1033080.
PR 19-APR-2001; 2001DE-1019294.
XX

PA (BOEH) BOEHRINGER INGELHEIM INT GMBH.

PI Schweizer N, Scherl-Mostageer M, Sommergruber W, Abseher R;
DR MPI: 2002-171704/22.
XX N-PSDB; ABA95507.
XX

PT New tumor-associated antigen B345, useful for diagnosis and
PT immunotherapy of tumors, also related nucleic acid and antibodies -
PS Claim 1; Page 85-88; 102pp; German.

CC This invention describes a novel tumour-associated antigen, designated
CC B345 which has cytostatic activity. B345 is involved in communication,
CC interaction and/or signal transduction with extracellular components and
CC ligands, especially in the metastatic potential of cancers, particularly
CC of the colon. B345 or its immunogenic fragments, also the DNA that
CC encodes it, are useful for immunotherapy of cancer, particularly
CC carcinoma of lung or colon. Antibodies raised against B345 are useful for
CC treatment and diagnosis of cancers that are associated with B345
CC expression, including their use for targeted delivery of cytotoxic or
CC radioactive agents. Probes derived from B345 can be used to detect
CC tumour-specific mutations in the B345 sequence, and can be used to screen
CC for B345 specific modulators. This sequence represents a human B345
CC tumour-associated antigen described in the invention.

SQ Sequence 836 AA;

Query Match 100.0%; Score 836; DB 23; Length 836;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 836; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALGLVLLGAARLPFGAFAFIALPRESNTITVLKGLPTLLAKCYIVI 60
DB 1 MAGLNCGVSIALGLVLLGAARLPFGAFAFIALPRESNTITVLKGLPTLLAKCYIVI 60
QY 61 SKRHITMLISGERIVFTFSCQSPENHFVIEIQKIDCMSPGCPFGVQLPSTSLPT 120
DB 61 SKRHITMLISGERIVFTFSCQSPENHFVIEIQKIDCMSPGCPFGVQLPSTSLPT 120
QY 121 LNRFTIMDKAKHSIGLELQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIMDKAKHSIGLELQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIRKQEGVKALHLPWFHPRNVSGFSIANSSIRLCTIESVFEGEGSATLMSANY 240
DB 181 GTVSRIRKQEGVKALHLPWFHPRNVSGFSIANSSIRLCTIESVFEGEGSATLMSANY 240
QY 241 PEGPPEDELMTMOFVPAHLRASVSFLNPNLSNCRKEEREVEYYPGSTTNPEVFKLEDK 300
DB 241 PEGPPEDELMTMOFVPAHLRASVSFLNPNLSNCRKEEREVEYYPGSTTNPEVFKLEDK 300
QY 301 QPGMAGNPNLSLQGCDDAOSPGILRLQFQVLYVHPONESNKIYVVDLSNERAMSLTIE 360
DB 301 QPGMAGNPNLSLQGCDDAOSPGILRLQFQVLYVHPONESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVKQSKKPVPGCVCLSERTCSNLTLTSGSKKISFLCDDLTRLMMANVEKTIISCTDH 420
DB 361 PRPVKQSKKPVPGCVCLSERTCSNLTLTSGSKKISFLCDDLTRLMMANVEKTIISCTDH 420
QY 421 RYCQRKSYSLQVPSDILHLPEVLHDFSMKLLVPRKORLSLVLPKQKLOQHTHEKPCNPSF 480
DB 421 RYCQRKSYSLQVPSDILHLPEVLHDFSMKLLVPRKORLSLVLPKQKLOQHTHEKPCNPSF 480
QY 481 SYLVASAIPSODLYFGSFCPGSGIKQIOVKONISVTLRTFAPSFQOEASROGLTVSFIPY 540
DB 481 SYLVASAIPSODLYFGSFCPGSGIKQIOVKONISVTLRTFAPSFQOEASROGLTVSFIPY 540
QY 541 FKESGVFTVPTDTSKVVYLRTPNMDRGLPSLTSVSNISVPRDOVACLTFKERSGVVQC 600
DB 541 FKESGVFTVPTDTSKVVYLRTPNMDRGLPSLTSVSNISVPRDOVACLTFKERSGVVQC 600
QY 601 TGRAPMTIQEORTAAEELFSLDEDEVLPRKSPFHHSFWNINISCSPTSGKQDLFLSVTLT 660
DB 601 TGRAPMTIQEORTAAEELFSLDEDEVLPRKSPFHHSFWNINISCSPTSGKQDLFLSVTLT 660
QY 661 PRYVDLTVILAAVGGVLLLSALGLIICVYKKKKKTKNGPAGVIYNGNINTEMPROPK 720
DB 661 PRYVDLTVILAAVGGVLLLSALGLIICVYKKKKKTKNGPAGVIYNGNINTEMPROPK 720
QY 721 KFOKGRKNDSHVYAVIBDVMYGHLLQDSSGSFLQPEVDVTRPFGTMGVCPPSPPTIC 780
DB 721 KFOKGRKNDSHVYAVIBDVMYGHLLQDSSGSFLQPEVDVTRPFGTMGVCPPSPPTIC 780
QY 781 SRAPTAKLATEEPPRSPSESESEPTTSHRNGVSKODDIPILANTQEPMEPAE 836
DB 781 SRAPTAKLATEEPPRSPSESESEPTTSHRNGVSKODDIPILANTQEPMEPAE 836

RESULT 3
AAM49640
ID AAM49640 standard; Protein; 749 AA.
XX AAM49640;
XX AC
XX 17-MAY-2002 (first entry)
XX DT
XX Human tumour-associated antigen B345 protein.
XX DE
XX Tumour-associated antigen; human; B345; cytosolic; cell communication;

KW cell interaction; signal transduction; metastasis; cancer; colon;
KW immunotherapy; carcinoma; lung; diagnosis.
XX
XX
OS Homo sapiens.
PN WO200204508-A1.
XX
XX 17-JAN-2002.
XX
XX 05-JUL-2001; 2001WO-EP0705.
XX
XX 07-JUL-2000; 2000DE-1033080.
PR 19-APR-2001; 2001DE-1019294.
XX
XX (BOEH) BOEHRINGER INGELHEIM INT GMBH.
PI Schaefer N, Scherl-Mostageer M, Sommergruber W, Abseher R;
XX WPI; 2002-171704/22.
DR N-PSDB; ABA99506.
XX
XX New tumor-associated antigen B345, useful for diagnosis and
PT immunotherapy of tumors, also related nucleic acid and antibodies -
XX
XX
PS Example 6; Page 76-78; 102pp; German.

This invention describes a novel tumour-associated antigen, designated B345 which has cytostatic activity. B345 is involved in communication, interaction and/or signal transduction with extracellular components and ligands, especially in the metastatic potential of cancers, particularly of the colon. B345 or its immunogenic fragments, also the DNA that encodes it, are useful for immunotherapy of cancer, particularly carcinoma of lung or colon. Antibodies raised against B345 are useful for treatment and diagnosis of cancers that are associated with B345 expression, including their use for targeted delivery of cytotoxic or radioactive agents. Probes derived from B345 can be used to detect tumour-specific mutations in the B345 sequence, and can be used to screen for B345 specific modulators. This sequence represents a human B345 tumour-associated antigen described in the invention.

XX
XX
SQ Sequence 749 AA;

Query Match 77.8%; Score 650; DB 23; Length 749;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 650; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 MLSTKSGRIVFTFSCQSPENHFVIEIQKIDCMSPGCPFGVQLPSTSLPTLNRFTFI 126
DB 1 MLSTKSGRIVFTFSCQSPENHFVIEIQKIDCMSPGCPFGVQLPSTSLPTLNRFTFI 126
QY 127 MDVKAHSIGLELQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTCSNGTVSRI 186
DB 127 MDVKAHSIGLELQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTCSNGTVSRI 186
QY 187 KMOEGVKALHLPWFHPRNVSGFSIANSSIRLCTIESVEGEGSATLMSANYPEGPPE 246
DB 187 KMOEGVKALHLPWFHPRNVSGFSIANSSIRLCTIESVEGEGSATLMSANYPEGPPE 246
QY 241 KMOEGVKALHLPWFHPRNVSGFSIANSSIRLCTIESVEGEGSATLMSANYPEGPPE 180
DB 241 KMOEGVKALHLPWFHPRNVSGFSIANSSIRLCTIESVEGEGSATLMSANYPEGPPE 180
QY 247 DELMTWQFVPAHLRASVSFLNPNLSNCRKEEREVEYYPGSTTNPEVFKLEDKQPGMA 306
DB 247 DELMTWQFVPAHLRASVSFLNPNLSNCRKEEREVEYYPGSTTNPEVFKLEDKQPGMA 306
QY 307 GNPNLSLQGCDDAOSPGILRLQFQVLYVHPONESNKIYVVDLSNERAMSLTIEPRPKQ 366
DB 307 GNPNLSLQGCDDAOSPGILRLQFQVLYVHPONESNKIYVVDLSNERAMSLTIEPRPKQ 366
QY 367 SRKFVPGCFVCLSERTCSNLTLTSGSKKISFLCDDLTRLMMANVEKTIISCTDHRVQCK 426
DB 367 SRKFVPGCFVCLSERTCSNLTLTSGSKKISFLCDDLTRLMMANVEKTIISCTDHRVQCK 426
QY 427 SYSLQVPSDILHLPEVLHDFSMKLLVPRKORLSLVLPKQKLOQHTHEKPCNPSFYLVA 486
DB 427 SYSLQVPSDILHLPEVLHDFSMKLLVPRKORLSLVLPKQKLOQHTHEKPCNPSFYLVA 486
QY 486 SYSLQVPSDILHLPEVLHDFSMKLLVPRKORLSLVLPKQKLOQHTHEKPCNPSFYLVA 420
DB 486 SYSLQVPSDILHLPEVLHDFSMKLLVPRKORLSLVLPKQKLOQHTHEKPCNPSFYLVA 420

Qy	48	AIISODLYFESFPCGSIKIQVKQNI SVTLRTFAPSQOASRGLTVSIPPEKEGV	546
Db	421	AIISQDLYFESFPCGSIKIQVKQNI SVTLRTFAPSQOASRGLTVSIPPEKEGV	480
Qy	547	FIVTPDTKSKVYLRITNMDRGLPSLTYSWNI SVPRDVA CLTFPKERSGVVCTGRAFM	606
Db	481	FIVTPDTKSKVYLRITNMDRGLPSLTYSWNI SVPRDVA CLTFPKERSGVVCTGRAFM	540
Qy	607	IIIEQRTAEIIFSLDEDLPKPSFHHSFWNI NSCSPTSGKQDLIFSVTLPRTVDL	666
Db	541	IIIEQRTAEIIFSLDEDLPKPSFHHSFWNI NSCSPTSGKQDLIFSVTLPRTVDL	600
Qy	667	TVLIIAAGGVLLLSALGLIICVKKKKKKNNKPAIGIYNGINTEMP 716	
Db	601	TVLIIAAGGVLLLSALGLIICVKKKKKKNNKPAIGIYNGINTEMP 650	
RESULT 4			
ID	AA91592	AA91592	
XX	AA91592	standard; Protein; 443 AA.	
AC	AA91592;		
XX			
DT	29-JUN-2000	(first entry)	
XX			
DE		Human secreted protein sequence encoded by gene 6 SEQ ID NO:265.	
XX			
KW		Human; secreted protein; diagnosis; cytostatic; immunosuppressive;	
KW		antiHIV; antiinflammatory; nootropic; neuroprotective; antiallergic;	
KW		osteopathic; antiaathritic; antibacterial; antidiabetic; antiasthma;	
KW		antiparasitic; cardiac; gene therapy; cancer; neurological disorder;	
KW		immune disease; inflammation; blood disorder; tumor.	
XX			
OS		Homo sapiens.	
XX			
PN	WO200006698-A1.		
XX			
PD	10-FEB-2000.		
XX			
PF	29-JUL-1999;	99WO-US17130.	
XX			
PR	30-JUL-1998;	98US-0094657.	
PR	05-AUG-1998;	98US-0095486.	
PR	06-AUG-1998;	98US-0095454.	
PR	06-AUG-1998;	98US-0095455.	
PR	12-AUG-1998;	98US-0096319.	
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
PI	Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;		
PI	Laflleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;		
PI	Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenksi M;		
DR	WPI; 2000-195282/17.		
XX			
PT		New isolated human genes and the secreted polypeptides they encode,	
PT		useful for diagnosis and treatment of e.g. cancers, neurological	
PT		disorders, immune diseases, inflammation or blood disorders -	
PS		Disclosure; Page 18; 634dp; English.	
XX			
CC		The polynucleotide sequences given in AAA26346 to AAA26458 encode the	
CC		human secreted proteins given in AA91451 to AA91691. The human secreted	
CC		proteins can have activities based on the tissues and cells they are	
CC		expressed in. Examples of the activities are: cytostatic;	
CC		immunosuppressive; antiHIV; antiinflammatory; nootropic; neuroprotective;	
CC		antiallergic; osteopathic; antiaathritic; antibacterial; antidiabetic;	
CC		antiasthma; antiparasitic; and cardiac. The polynucleotides and their	
CC		corresponding secreted proteins are useful for preventing, treating or	
CC		ameliorating medical conditions, e.g. by protein or gene therapy. Also	
CC		pathological conditions can be diagnosed by determining the amount of the	
CC		proteins in a sample or by determining the presence of mutations in the	

CC		polynucleotides. Specific uses are described for each of the
CC		polynucleotides, based on which tissues they are most highly expressed
CC		in, and include developing products for the diagnosis or treatment of
CC		cancer, tumors, neurodegenerative disorders, developmental abnormalities
CC		and foetal deficiencies, blood disorders, diseases of the immune system,
CC		autoimmune diseases, hepatic and renal disease, inflammation,
CC		allergies, Alzheimer's and behavioural disorders, schizophrenia,
CC		osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
CC		transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
CC		cardiovascular disorders, reproductive disorders, gastrointestinal
CC		disorders, respiratory disorders and metabolic disorders. The proteins
CC		or polynucleotides can also be used as food additives or preservatives.
CC		The proteins are also useful for identifying their binding partners.
CC		AA2637 to AA26345 and AA91450 are sequences used in the
CC		exemplification of the present invention.
XX		
SQ	Sequence	443 AA;
	Query Match	49.4%; Score 413; DB 21; Length 443;
	Best Local Similarity	100.0%; Pred. No. 0;
	Matches 413;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1	MAGINCVSIALGLVLLGGAARLPARGAEPALPRESNITVLTKGTPTLLAKPCYIVI 60
DB	30	MAGINCVSIALGLVLLGGAARLPARGAEPALPRESNITVLTKGTPTLLAKPCYIVI 89
QY	61	SKRIITMLSTKSGRIVFTFSCSPENHFVEIOKNIDCMGSCPFGGVOLQPSTSLIPT 120
DB	90	SKRIITMLSTKSGRIVFTFSCSPENHFVEIOKNIDCMGSCPFGGVOLQPSTSLIPT 149
QY	121	LNRFFIMDVVAHKSIGLELOPSTRRLQIIGGESCPGVTHSISGRIDATVVRIGTFCSN 180
DB	150	LNRFFIMDVVAHKSIGLELOPSTRRLQIIGGESCPGVTHSISGRIDATVVRIGTFCSN 209
QY	181	GTVSRIRIKMGESVKMALHPMFHPNPVNGFSIANSSIKRLCTIIISVFEGESATLSMANY 240
DB	210	GTVSRIRIKMGESVKMALHPMFHPNPVNGFSIANSSIKRLCTIIISVFEGESATLSMANY 269
QY	241	PEGFEDELMTWQFVVPAHLRASVFLNFNLNCERKEERYEYIIPGSTTPEVFKLEDK 300
DB	270	PEGFEDELMTWQFVVPAHLRASVFLNFNLNCERKEERYEYIIPGSTTPEVFKLEDK 329
QY	301	QPGMAINFPLSLGGCPDADOSPGLIELQOVLYOHNPONESNKTYVDLNSRAMSLITE 360
DB	330	QPGMAINFPLSLGGCPDADOSPGLIELQOVLYOHNPONESNKTYVDLNSRAMSLITE 389
QY	361	PRPVKGRRKFVPGCFVCLESRTCSNLTTLNCGSHKHSIFLCDDLTRLMMANVEK 413
DB	390	PRPVKGRRKFVPGCFVCLESRTCSNLTTLNCGSHKHSIFLCDDLTRLMMANVEK 442
RESULT 5		
AAB95558	ID	AAB95558 standard; Protein: 649 AA.
XX		
AC	AAB95558;	
XX		
D7	26-JUN-2001	(first entry)
XX		
DE		Human protein sequence SEQ ID NO:18191.
XX		
KW		Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX		
OS	Homo sapiens.	
XX	EPI074617-A2.	
PN		
XX		
PD	07-FEB-2001.	
XX		
PF	28-JUL-2000; 2000EP-0116126.	
XX		
JU	29-JUL-1999; 99JP-0248036.	
R	27-AUG-1999; 99JP-0300253.	
RR		

```

PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 18191; 2537bp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC polynucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH1628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 649 AA;
Query Match 45.7%; Score 382; DB 22; Length 649;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 582; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 253 QFVVPAAHLRASVSFLNPNLSNCRKEEVEYYIPGSTNPEVFKLEDKQPGMAGNFNLS 312
DB 66 QFVVPAAHLRASVSFLNPNLSNCRKEEVEYYIPGSTNPEVFKLEDKQPGMAGNFNLS 125
OY 313 LOGCDODQSPGILRLQFOVLVQHONSNKTYVVDLSNERAMSTTTPRPVKQSRKVP 372
DB 126 LOGCDODQSPGILRLQFOVLVQHONSNKTYVVDLSNERAMSTTTPRPVKQSRKVP 185
OY 373 GCFVCLERTSSNLTLTSGSKHKISFLCDDTLRLMMVVEKTICTDHRVCORKSYSLOV 432
DB 186 GCFVCLERTSSNLTLTSGSKHKISFLCDDTLRLMMVVEKTICTDHRVCORKSYSLOV 245
OY 433 PSDILHLPVELHDFESKLLVLPKDLRLVLPKQLQOHTHEKPCNTSPSYLVASAPSD 492
DB 246 PGDILHLPVELHDFESKLLVLPKDLRLVLPKQLQOHTHEKPCNTSPSYLVASAPSD 305
OY 493 LYFGSFCGGSISIKQIQVONISVTLRTAPSFQASAGLTVSTIPFKKEGVTTVPD 552
DB 306 LYFGSFCGGSISIKQIQVONISVTLRTAPSFQASAGLTVSTIPFKKEGVTTVPD 365
OY 553 TKSRYLATPMMDDGLPSLTYSVNNISVPRDOVACLTFKERSGVVCTGAFMIIOBOR 612
DB 366 TKSRYLATPMMDDGLPSLTYSVNNISVPRDOVACLTFKERSGVVCTGAFMIIOBOR 425
OY 613 TRAEIIFSLDSDVLPKPSFHHHSFWNINSNCSPTSGKQDLDFSVTLTPRTVDLTVIIA 672
DB 425 TRAEIIFSLDSDVLPKPSFHHHSFWNINSNCSPTSGKQDLDFSVTLTPRTVDLTVIIA 672

```

```

DB 426 TRAEIIFSLDSDVLPKPSFHHHSFWNINSNCSPTSGKQDLDFSVTLTPRTVDLTVIIA 485
OY 673 AVGGGVLILLSALGILICCVKKKKKTKGPAVGILYNGNINTEMROPKPKQGRKNDNSH 732
DB 486 AVGGGVLILLSALGILICCVKKKKKTKGPAVGILYNGNINTEMROPKPKQGRKNDNSH 545
OY 733 VYAVIEDTMVYGHLLDSSGSLQPEVDYTRPFQGTWGVCPSPPTICSRAPYAKLATEE 792
DB 546 VYAVIEDTMVYGHLLDSSGSLQPEVDYTRPFQGTWGVCPSPPTICSRAPYAKLATEE 605
OY 793 PPRSPPSESEPTFSHPNNGDVSSKOTDIPLNTQEPMEPAE 836
DB 606 PPRSPPSESEPTFSHPNNGDVSSKOTDIPLNTQEPMEPAE 649
RESULT 6
AAU12252
ID AAU12252 standard; Protein: 343 AA.
XX
AC AAU12252;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human PRO5773 polypeptide sequence.
XX
KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;
KW breast; prostate; cervical; tumor necrosis factor-alpha; TNF-alpha;
KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
KW adipocyte; A-peptide; factor VIIa; gene therapy.
XX
OS Homo sapiens.
XX
MN MO200140466-A2.
PD 07-JUN-2001.
XX
PF 01-DEC-2000; 2000WO-US32678.
XX
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 09-DEC-1999; 99US-U170262.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 30-DEC-1999; 99WO-US31243.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 20-MAR-2000; 2000WO-US07377.
PR 21-MAR-2000; 2000WO-US07532.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 10-NOV-2000; 2000WO-US30873.
PA (GETH ) GENENTECH INC.
XX
PI Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerlitsen ME, Goddard A, Godowski PJ, Guirney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI; 2001-408281/43.

```


DR N-Peptide; AAS21324.
XX Isolated, secretory and transmembrane PRO polypeptide used to detect
PT other PRO polypeptides, link bioactive molecules to cells expressing
PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
PT lung, breast, prostate, cervical -
XX Claim 12; Fig 162; 813pp; English.
XX
XX AAU12172-AAU12446 represent novel human secretory and transmembrane
CC PRO polypeptides. The PRO polypeptides are useful to detect other
CC PRO polypeptides, to link bioactive molecules to cells expressing
CC PRO polypeptides, to modulate biological activities of cells expressing
CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
CC polypeptide expression in a cell sample to that in a control sample.
CC Some of the 275 sequences are also useful to stimulate the release of
CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
CC proliferation or differentiation of chondrocytes, the proliferation or
CC gene expression in pericyte cells, the release of proteoglycans from
CC cartilage, the proliferation of inner ear utricular supporting cells or
CC of T-lymphocytes, the release of a cytokine from peripheral blood
CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of
CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
CC skeletal muscle cells or by adipocytes, or inhibit binding of A-peptide
CC to factor VIIa. The PRO polypeptides can be used in assays to identify
CC molecules involved in binding interactions. The polynucleotides encoding
CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
CC transgenic or knock out animals and can be used in gene therapy.
XX
XX Sequence 343 AA:

Query Match 40.8%; Score 341; DB 22; Length 343;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLGAAALPRGAEPFIALPRESNITVILIKIGTPTLLAKCYIYI 60
DB 1 MAGLNCGVSIALLGVLLGAAALPRGAEPFIALPRESNITVILIKIGTPTLLAKCYIYI 60
QY 61 SRKHITMISIKSGERIVFTFSCSPENHVIETIKKIDMSGCPGEGVQLPSTSLPT 120
DB 61 SRKHITMISIKSGERIVFTFSCSPENHVIETIKKIDMSGCPGEGVQLPSTSLPT 120
QY 121 LNRFTIMDKAKHSIGLELOFSIPRLRQIPGSSCPDGYTHSISGRIDATVIRIGTFCSN 180
DB 121 LNRFTIMDKAKHSIGLELOFSIPRLRQIPGSSCPDGYTHSISGRIDATVIRIGTFCSN 180
QY 181 GTVSRIKQOEGVYKALHLFPFHPRNVSIGFSIANRSSIKRLCTIESVFEBSGATLMSANY 240
DB 181 GTVSRIKQOEGVYKALHLFPFHPRNVSIGFSIANRSSIKRLCTIESVFEBSGATLMSANY 240
QY 241 PGSPPEDELMTQFVPAHLRASVSFLNLSNCKEKERVYIYI PGSTTNEVFLBEX 300
DB 241 PGSPPEDELMTQFVPAHLRASVSFLNLSNCKEKERVYIYI PGSTTNEVFLBEX 300
QY 301 QPGNAGNFNLSLQGCDDQASPGILRLQFVLYVHPQNES 341
DB 301 QPGNAGNFNLSLQGCDDQASPGILRLQFVLYVHPQNES 341

RESULT 7
ABU66650
ID ABU66650 standard; Protein; 343 AA.

XX ABU66650;

XX 23-MAY-2003 (first entry)

XX Human PRO polypeptide #81.

XX Human, PRO polypeptide; secreted and transmembrane protein;
KW tumour necrosis factor-alpha; TNF-alpha; blood; proliferation;

KW differentiation; chondrocyte; tumour; genetic disorder;
KW cytostatic.
XX
XX Homo sapiens.
OS
XX US2003036180-A1.
PN
XX 20-FEB-2003.
PD
XX
XX 09-MAY-2002; 2002US-0143114.
PF
XX 31-MAR-1997; 97WO-US05230.
XX 12-JUN-1998; 98WO-US12456.
XX 14-JUL-1998; 98WO-US14552.
PR 28-AUG-1998; 98WO-US17888.
PR 10-SEP-1998; 98WO-US18824.
PR 14-SEP-1998; 98WO-US19093.
PR 14-SEP-1998; 98WO-US19094.
PR 14-SEP-1998; 98WO-US19177.
PR 16-SEP-1998; 98WO-US19330.
PR 17-SEP-1998; 98WO-US19437.
PR 07-OCT-1998; 98WO-US21141.
PR 29-OCT-1998; 98WO-US22991.
PR 20-NOV-1998; 98WO-US22992.
PR 01-DEC-1998; 98WO-US24855.
PR 05-JAN-1999; 98WO-US25106.
PR 08-MAR-1999; 98WO-US05028.
PR 10-MAR-1999; 98WO-US05190.
PR 20-APR-1999; 98WO-US08615.
PR 14-MAY-1999; 98WO-US10733.
PR 02-JUN-1999; 98WO-US12252.
PR 01-SEP-1999; 98WO-US20111.
PR 08-SEP-1999; 98WO-US20594.
PR 13-SEP-1999; 98WO-US20944.
PR 15-SEP-1999; 98WO-US21090.
PR 15-SEP-1999; 98WO-US21547.
PR 05-OCT-1999; 98WO-US23089.
PR 29-NOV-1999; 98WO-US28214.
PR 30-NOV-1999; 98WO-US28313.
PR 30-NOV-1999; 98WO-US28409.
PR 01-DEC-1999; 98WO-US28301.
PR 01-DEC-1999; 98WO-US28634.
PR 02-DEC-1999; 98WO-US28551.
PR 02-DEC-1999; 98WO-US28564.
PR 16-DEC-1999; 98WO-US28565.
PR 20-DEC-1999; 98WO-US30095.
PR 20-DEC-1999; 98WO-US30911.
PR 22-DEC-1999; 98WO-US30999.
PR 30-DEC-1999; 98WO-US30720.
PR 30-DEC-1999; 98WO-US31243.
PR 05-JAN-2000; 98WO-US31274.
PR 06-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00277.
PR 11-FEB-2000; 2000WO-US00376.
PR 18-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 22-FEB-2000; 2000WO-US04342.
PR 24-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05746.
PR 10-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06319.
PR 20-MAR-2000; 2000WO-US06884.
PR 21-MAR-2000; 2000WO-US07377.
PR 30-MAR-2000; 2000WO-US07532.
PR 17-MAY-2000; 2000WO-US08439.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.

28-JUL-2000; 2000WO-US20710.
 PR 11-AUG-2000; 2000WO-US22031.
 PR 23-AUG-2000; 2000WO-US23252.
 PR 24-AUG-2000; 2000WO-US23328.
 PR 08-NOV-2000; 2000WO-US30952.
 PR 10-NOV-2000; 2000WO-US30873.
 PR 01-DEC-2000; 2000WO-US34678.
 PR 20-DEC-2000; 2000WO-US34956.
 PR 28-FEB-2001; 2001WO-US06520.
 PR 01-MAR-2001; 2001WO-US06666.
 PR 25-MAY-2001; 2001WO-US17092.
 PR 01-JUN-2001; 2001WO-US17800.
 PR 20-JUN-2001; 2001WO-US19692.
 PR 22-JUN-2001; 2001WO-US20116.
 PR 29-JUN-2001; 2001WO-US21066.
 PR 09-JUL-2001; 2001WO-US21735.
 PR 20-DEC-2000; 2000US-0747259.
 PR 28-FEB-2001; 2001US-0796498.
 PR 09-MAR-2001; 2001US-0802706.
 PR 14-MAR-2001; 2001US-0808689.
 PR 22-MAR-2001; 2001US-0816744.
 PR 05-APR-2001; 2001US-0828366.
 PR 10-MAY-2001; 2001US-0854208.
 PR 10-MAY-2001; 2001US-0854280.
 PR 18-MAY-2001; 2001US-0860216.
 PR 25-MAY-2001; 2001US-0866028.
 PR 01-JUN-2001; 2001US-0872035.
 PR 05-JUN-2001; 2001US-0874503.
 PR 14-JUN-2001; 2001US-0882636.
 PR 19-JUN-2001; 2001US-0886342.
 PR 21-JUN-2001; 2001US-0887879.
 PR 18-JUL-2001; 2001US-0908827.
 PR 06-AUG-2001; 2001US-0924419.
 PR 09-AUG-2001; 2001US-0927796.
 PR 16-AUG-2001; 2001US-0931836.
 PR 19-DEC-2001; 2001US-0028072.
 (GETH) GENENTECH INC.
 Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W, Gerltien ME, Goddard A, Godowski PJ, Gurney AU, Sherwood S, Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z; MPI: 2003-332040/31.
 DR N-PSDB; ACN035683.
 New secreted and transmembrane PRO nucleic acids, useful for gene therapy, in chromosome and gene mapping, as chromosome markers, in tissue typing, and in chromosome identification -
 Claim 12, Fig 162; 660p; English.

XX SQ Sequence 343 AA;
 Query Match 40.8%; Score 341; DB 24; Length 343;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAGINCVSIALLVLLGAARLRGAFAEIALPRESNTIVLTKGTPTLLARPCYIVI 60
 DB 1 MAGINCVSIALLVLLGAARLRGAFAEIALPRESNTIVLTKGTPTLLARPCYIVI 60
 QY 61 SKRIITMISTKSGRIYFTFSCSPENHFVIEIQKNDICMGPCPFGEVQLOPSTSLPT 120
 DB 61 SKRIITMISTKSGRIYFTFSCSPENHFVIEIQKNDICMGPCPFGEVQLOPSTSLPT 120
 QY 121 LNRFTIMVVKAKHSGIGLEQPSIRLQIGPESCPDGVHSIGRIDATVVRIGTFCPSN 180
 DB 121 LNRFTIMVVKAKHSGIGLEQPSIRLQIGPESCPDGVHSIGRIDATVVRIGTFCPSN 180
 QY 181 GTVSRIRKQEGVKMALHLPWHPNNVSGFSIANSSIRKLCIISVPEGSGATLMSANY 240
 DB 181 GTVSRIRKQEGVKMALHLPWHPNNVSGFSIANSSIRKLCIISVPEGSGATLMSANY 240
 QY 241 PEGFPEDELMTWQVVPVPAHLPAVSPLNFNISNCRKEERYEYIIPGSTNPEVFKLEDK 300
 DB 241 PEGFPEDELMTWQVVPVPAHLPAVSPLNFNISNCRKEERYEYIIPGSTNPEVFKLEDK 300
 QY 301 QPGNMAGNFNLSLQCCDDAOSPGLRLQFOVLVQHPONES 341
 DB 301 QPGNMAGNFNLSLQCCDDAOSPGLRLQFOVLVQHPONES 341
 RESULT 8
 ABU66926
 ID ABU66926 standard; Protein; 343 AA.
 XX ABU66926;
 AC
 XX
 DT 27-MAY-2003 (first entry)
 XX
 DE Human secreted/transmembrane, PRO, protein SEQ ID 162.
 XX
 DE Human; secreted protein; transmembrane protein; PRO;
 XX inflammatory disease; organ failure; atherosclerosis; cardiac injury;
 XX infertility; birth defects; premature aging; AIDS; biosensor;
 XX acquired immunodeficiency syndrome; cancer; diabetic complication;
 XX bioreactor; tumour.
 OS Homo sapiens.
 XX
 PN US2003032155-A1.
 XX
 PD 13-FEB-2003.
 XX
 PD 03-MAY-2002; 2002US-0137865.
 XX
 PF 31-MAR-1997; 97WO-US05230.
 XX 12-JUN-1998; 98WO-US12456.
 XX 14-JUL-1998; 98WO-US14552.
 XX 28-AUG-1998; 98WO-US17888.
 XX 10-SEP-1998; 98WO-US18824.
 XX 14-SEP-1998; 98WO-US19093.
 XX 14-SEP-1998; 98WO-US19094.
 XX 14-SEP-1998; 98WO-US19177.
 XX 16-SEP-1998; 98WO-US19330.
 XX 17-SEP-1998; 98WO-US19437.
 XX 07-OCT-1998; 98WO-US21141.
 XX 29-OCT-1998; 98WO-US22991.
 XX 29-OCT-1998; 98WO-US22992.
 XX 20-NOV-1998; 98WO-US24855.
 XX 01-DEC-1998; 98WO-US25108.
 XX 05-JAN-1999; 99WO-US00106.
 XX 08-MAR-1999; 99WO-US05028.

PR 10-MAR-1999; 99MO-US05190.
PR 20-APR-1999; 99MO-US08615.
PR 14-MAY-1999; 99MO-US10733.
PR 02-JUN-1999; 99MO-US12252.
PR 01-SEP-1999; 99MO-US20111.
PR 08-SEP-1999; 99MO-US20594.
PR 13-SEP-1999; 99MO-US20944.
PR 15-SEP-1999; 99MO-US21090.
PR 15-SEP-1999; 99MO-US21547.
PR 05-OCT-1999; 99MO-US21089.
PR 29-NOV-1999; 99MO-US28214.
PR 30-NOV-1999; 99MO-US28313.
PR 30-NOV-1999; 99MO-US28409.
PR 01-DEC-1999; 99MO-US28301.
PR 01-DEC-1999; 99MO-US28634.
PR 02-DEC-1999; 99MO-US28551.
PR 02-DEC-1999; 99MO-US28564.
PR 02-DEC-1999; 99MO-US28565.
PR 16-DEC-1999; 99MO-US30095.
PR 20-DEC-1999; 99MO-US30911.
PR 22-DEC-1999; 99MO-US30999.
PR 30-DEC-1999; 99MO-US31243.
PR 30-DEC-1999; 99MO-US31274.
PR 05-JAN-2000; 2000MO-US0219.
PR 06-JAN-2000; 2000MO-US0277.
PR 06-JAN-2000; 2000MO-US0376.
PR 11-FEB-2000; 2000MO-US03565.
PR 18-FEB-2000; 2000MO-US04341.
PR 18-FEB-2000; 2000MO-US04342.
PR 22-FEB-2000; 2000MO-US04414.
PR 24-FEB-2000; 2000MO-US04914.
PR 24-FEB-2000; 2000MO-US05004.
PR 01-MAR-2000; 2000MO-US05601.
PR 02-MAR-2000; 2000MO-US05746.
PR 02-MAR-2000; 2000MO-US05841.
PR 10-MAR-2000; 2000MO-US06119.
PR 15-MAR-2000; 2000MO-US06884.
PR 20-MAR-2000; 2000MO-US07377.
PR 21-MAR-2000; 2000MO-US07532.
PR 30-MAR-2000; 2000MO-US08439.
PR 17-MAY-2000; 2000MO-US13705.
PR 22-MAY-2000; 2000MO-US14042.
PR 30-MAY-2000; 2000MO-US14941.
PR 02-JUN-2000; 2000MO-US15264.
PR 28-JUL-2000; 2000MO-US20710.
PR 11-AUG-2000; 2000MO-US22031.
PR 23-AUG-2000; 2000MO-US23522.
PR 24-AUG-2000; 2000MO-US23528.
PR 08-NOV-2000; 2000MO-US30952.
PR 10-NOV-2000; 2000MO-US30873.
PR 01-DEC-2000; 2000MO-US32678.
PR 20-DEC-2000; 2000MO-US34956.
PR 28-FEB-2001; 2001MO-US06520.
PR 01-MAR-2001; 2001MO-US06566.
PR 25-MAY-2001; 2001MO-US17092.
PR 01-JUN-2001; 2001MO-US17800.
PR 20-JUN-2001; 2001MO-US19692.
PR 22-JUN-2001; 2001MO-US20116.
PR 29-JUN-2001; 2001MO-US21066.
PR 09-JUL-2001; 2001MO-US21735.
PR 20-DEC-2000; 2000MO-US21735.
PR 28-FEB-2001; 2001MO-US21735.
PR 09-MAR-2001; 2001MO-US21735.
PR 14-MAR-2001; 2001MO-US21735.
PR 22-MAR-2001; 2001MO-US21735.
PR 05-APR-2001; 2001MO-US21735.
PR 10-MAY-2001; 2001MO-US21735.
PR 18-MAY-2001; 2001MO-US21735.
PR 25-MAY-2001; 2001MO-US21735.
PR 01-JUN-2001; 2001MO-US21735.

PR 05-JUN-2001; 2001US-0874503.
PR 14-JUN-2001; 2001US-0882636.
PR 19-JUN-2001; 2001US-0886342.
PR 21-JUN-2001; 2001US-0887879.
PR 18-JUL-2001; 2001US-0908827.
PR 06-AUG-2001; 2001US-0924419.
PR 09-AUG-2001; 2001US-0927796.
PR 16-AUG-2001; 2001US-0931836.
PR 19-DEC-2001; 2001US-0028072.
XX
XX (GENE) GENENTECH INC.
XX
XX Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W,
XX Gerritsen ME, Goddard A, Godowski PJ, Gunney AL, Sherwood S,
XX Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX
XX MPI: 2003-331925/31.
XX N-PSDB; ACN04104.
XX
XX New secreted and transmembrane nucleic acids and polypeptides,
XX designated as PRO, useful for treating inflammation, organ failure,
XX atherosclerosis, cardiac injury, infertility, birth defects, premature
XX aging, AIDS, or cancer -
XX
XX Claim 12; Fig 162; 659pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising, or which is
XX at least 80% identical to, or the full-length coding sequence of, any of
XX the 275 nucleotide sequences, encoding the corresponding PRO polypeptide
XX (one of 275 secreted or transmembrane proteins). The nucleic acid
XX further comprises the full-length coding sequence of the DNA deposited
XX under American Type Culture Collection (ATCC) accession number in a list
XX given in the specification. Also included are vectors and host
XX cells for producing PRO proteins, PRO fusion proteins, anti-PRO
XX antibodies, PRO extracellular domains and mature sequences, methods
XX of detecting PRO proteins, methods for stimulating the release of
XX TNF-alpha (tumour necrosis factor alpha) from human blood.
XX (and the proliferation of differentiation of chondrocyte cells, the
XX proliferation of, or gene expression in pericyte cells, the release or
XX proteoglycans from cartilage, proliferation of inner ear utricular
XX supporting cells, the proliferation of T-lymphocyte cells, the release
XX of a cytokine from peripheral blood mononuclear cells (PBMC), or the
XX proliferation of endothelial cells), a method for modulating the uptake
XX of glucose or free fatty acid (FFA) by skeletal muscle cells,
XX a method for inhibiting the binding of A-peptide to factor VIIA,
XX or the differentiation of adipocyte cells, a method for detecting the
XX presence of a tumour in a mammal and an oligonucleotide probe derived
XX from any of the nucleotide sequences cited above. The nucleic acids and
XX polypeptides are useful for treating inflammatory diseases, organ
XX failure, atherosclerosis, cardiac injury, infertility, birth defects,
XX premature aging, AIDS (acquired immunodeficiency syndrome), cancer, or
XX diabetic complications. The nucleic acids are useful as hybridisation
XX probes, in chromosome and gene mapping, and in generating antisense RNA
XX or DNA. The polypeptides are useful as pharmaceuticals, diagnostics,
XX bioassays or bioreactors. Both are useful in tissue typing.
XX The present sequence represents a PRO protein of the invention.
XX
SQ Sequence 343 AA:
Query Match 40.8%; Score 341; DB 24; Length 343;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAGLNGVSIALGLVLLGAARLPAGAFETALPRESNITVLYKGTPTLAKPCYVI 60
Db 1 MAGLNGVSIALGLVLLGAARLPAGAFETALPRESNITVLYKGTPTLAKPCYVI 60
QY 61 SKRHITMLSIKSGERIVTFSCQSPENHFVLEIQKINDMSGPCFGEVQLPSTSLPT 120
Db 61 SKRHITMLSIKSGERIVTFSCQSPENHFVLEIQKINDMSGPCFGEVQLPSTSLPT 120
QY 121 LNRFTIWDVKAHKSIGLELQPSIRPLRQIGPESCPDGVTHSISRIDATVVRIGTFCSN 180

Db 121 LNRTITWVKAHKSIGLEIQFIRPLRQIGPESCPDGTHTISGRIDATVTRIGTFCSN 180
 Qy 181 GTVSRIRKQEGVKMLHLFWPFRPNVSGFSIANRSSIKRLCTIESEVFESEGSATLMSANY 240
 Db 181 GTVSRIRKQEGVKMLHLFWPFRPNVSGFSIANRSSIKRLCTIESEVFESEGSATLMSANY 240
 Qy 241 PEGFPEDEIMTQFVVPALHRSVSGFLNLSNCRKEERVEYIPGTTNDEVEFKLEDK 300
 Db 241 PEGFPEDEIMTQFVVPALHRSVSGFLNLSNCRKEERVEYIPGTTNDEVEFKLEDK 300
 Qy 301 QPGNMAGNPNLSLQCGDDOASPGILRLQFQVLNQHPONES 341
 Db 301 QPGNMAGNPNLSLQCGDDOASPGILRLQFQVLNQHPONES 341
 RESULT 9
 ABUS9731
 ID ABUS9731 standard; Protein; 343 AA.
 XX
 AC ABUS9731;
 XX
 DT 13-MAY-2003 (first entry)
 XX
 DE Novel secreted and transmembrane protein PRO5773.
 XX
 KW Human; PRO; hypertrophy of neonatal heart; angiogenesis; wound healing;
 KW cardiac insufficiency disorder; cancer; tumour; immune response;
 KW adrenal cortical capillary endothelial growth; c-fos induction;
 KW vascular endothelial growth factor inhibition; VEGF inhibition;
 KW endothelial cell growth inhibitor; T-lymphocytes stimulation;
 KW retinal neurons cell survival; rod photoreceptor cell survival;
 KW retinal disorder; retinitis pigmentosa; kidney disorder;
 KW mammalian kidney mesangial cell proliferation; Berger disease;
 KW dermatitis; herpeticiformis; Cronin's disease; chondrocyte proliferation;
 KW chondrocyte redifferentiation; sports injury; arthritis.
 XX
 OS Homo sapiens.
 XX
 PN US2003017563-A1.
 XX
 PD 23-JAN-2003.
 XX
 PF 07-MAY-2002; 2002US-0140808.
 XX
 XX 31-MAR-1997; 98WO-US05230.
 PR 12-JUN-1998; 98WO-US12456.
 PR 14-JUL-1998; 98WO-US14552.
 PR 28-AUG-1998; 98WO-US17888.
 PR 10-SEP-1998; 98WO-US18824.
 PR 14-SEP-1998; 98WO-US19093.
 PR 14-SEP-1998; 98WO-US19094.
 PR 14-SEP-1998; 98WO-US19177.
 PR 16-SEP-1998; 98WO-US19330.
 PR 17-SEP-1998; 98WO-US19437.
 PR 07-OCT-1998; 98WO-US21141.
 PR 29-OCT-1998; 98WO-US22991.
 PR 29-OCT-1998; 98WO-US22992.
 PR 20-NOV-1998; 98WO-US24855.
 PR 01-DEC-1998; 98WO-US25108.
 PR 05-JAN-1999; 99WO-US00106.
 PR 08-MAR-1999; 99WO-US05028.
 PR 10-MAR-1999; 99WO-US05190.
 PR 20-APR-1999; 99WO-US08615.
 PR 14-MAY-1999; 99WO-US10733.
 PR 02-JUN-1999; 99WO-US12252.
 PR 01-SEP-1999; 99WO-US20111.
 PR 08-SEP-1999; 99WO-US20594.
 PR 13-SEP-1999; 99WO-US20944.
 PR 15-SEP-1999; 99WO-US21090.
 PR 15-SEP-1999; 99WO-US21547.
 PR 05-OCT-1999; 99WO-US23089.
 PR 29-NOV-1999; 99WO-US28214.
 PR 30-NOV-1999; 99WO-US28313.

PR 30-NOV-1999; 99WO-US28409.
 PR 01-DEC-1999; 99WO-US28301.
 PR 02-DEC-1999; 99WO-US28634.
 PR 02-DEC-1999; 99WO-US28551.
 PR 02-DEC-1999; 99WO-US28564.
 PR 02-DEC-1999; 99WO-US28565.
 PR 16-DEC-1999; 99WO-US30095.
 PR 20-DEC-1999; 99WO-US30911.
 PR 22-DEC-1999; 99WO-US30999.
 PR 30-DEC-1999; 99WO-US30720.
 PR 30-DEC-1999; 99WO-US31243.
 PR 05-JAN-2000; 99WO-US31274.
 PR 06-JAN-2000; 2000WO-US00219.
 PR 06-JAN-2000; 2000WO-US00277.
 PR 11-FEB-2000; 2000WO-US00376.
 PR 18-FEB-2000; 2000WO-US03565.
 PR 18-FEB-2000; 2000WO-US04341.
 PR 18-FEB-2000; 2000WO-US04342.
 PR 22-FEB-2000; 2000WO-US04914.
 PR 24-FEB-2000; 2000WO-US04914.
 PR 24-FEB-2000; 2000WO-US04914.
 PR 01-MAR-2000; 2000WO-US05004.
 PR 01-MAR-2000; 2000WO-US05601.
 PR 02-MAR-2000; 2000WO-US05746.
 PR 02-MAR-2000; 2000WO-US05841.
 PR 10-MAR-2000; 2000WO-US06319.
 PR 15-MAR-2000; 2000WO-US06884.
 PR 20-MAR-2000; 2000WO-US07377.
 PR 21-MAR-2000; 2000WO-US07537.
 PR 30-MAR-2000; 2000WO-US08439.
 PR 17-MAY-2000; 2000WO-US13705.
 PR 22-MAY-2000; 2000WO-US14042.
 PR 30-MAY-2000; 2000WO-US14941.
 PR 02-JUN-2000; 2000WO-US15264.
 PR 28-JUN-2000; 2000WO-US20710.
 PR 11-AUG-2000; 2000WO-US22031.
 PR 23-AUG-2000; 2000WO-US23522.
 PR 24-AUG-2000; 2000WO-US23328.
 PR 08-NOV-2000; 2000WO-US30952.
 PR 10-NOV-2000; 2000WO-US30873.
 PR 01-DEC-2000; 2000WO-US32678.
 PR 20-DEC-2000; 2000WO-US34956.
 PR 28-FEB-2001; 2001WO-US06520.
 PR 01-MAR-2001; 2001WO-US06666.
 PR 25-MAY-2001; 2001WO-US17092.
 PR 01-JUN-2001; 2001WO-US17800.
 PR 20-JUN-2001; 2001WO-US19692.
 PR 22-JUN-2001; 2001WO-US20116.
 PR 29-JUN-2001; 2001WO-US21066.
 PR 09-JUL-2001; 2001WO-US21735.
 PR 20-DEC-2000; 2000US-0747259.
 PR 28-FEB-2001; 2001US-0796498.
 PR 09-MAR-2001; 2001US-0802706.
 PR 14-MAR-2001; 2001US-0806889.
 PR 22-MAR-2001; 2001US-0816744.
 PR 05-APR-2001; 2001US-0828366.
 PR 10-MAY-2001; 2001US-0854208.
 PR 10-MAY-2001; 2001US-0854280.
 PR 18-MAY-2001; 2001US-0860216.
 PR 25-MAY-2001; 2001US-0866028.
 PR 01-JUN-2001; 2001US-0866034.
 PR 05-JUN-2001; 2001US-0872035.
 PR 14-JUN-2001; 2001US-0874503.
 PR 14-JUN-2001; 2001US-0882636.
 PR 19-JUN-2001; 2001US-0886342.
 PR 21-JUN-2001; 2001US-0887879.
 PR 18-JUL-2001; 2001US-0908827.
 PR 06-AUG-2001; 2001US-0924419.
 PR 09-AUG-2001; 2001US-0927796.
 PR 16-AUG-2001; 2001US-0931836.
 PR 19-DEC-2001; 2001US-0028072.
 XX
 XX
 PA (GETH) GENENTECH INC.
 XX

PI Baker KP, Beresini M, DeForge L, Deenoyers L, Filyroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX WPI: 2003-148238/14.
XX N-PSDB; ABR89221.
DR
XX
XX Novel isolated PRO polypeptides e.g., PRO826, PRO1068, PRO1184, PRO1346
PT and PRO1375, which stimulate proliferation of stimulated T-lymphocytes
PT are therapeutically useful for enhancing immune response and in cancer
PT treatments -
XX
XX Claim 12: Fig 162; 659pp; English.
XX
XX The invention describes an isolated human PRO polypeptide. The PRO
CC polypeptides are useful in detecting PRO polypeptides in a sample, in
CC linking a bioactive molecule to a cell expressing a PRO polypeptide, and
CC in modulating at least one biological activity of a cell expressing a PRO
CC polypeptide. PRO1312 stimulates hypertrophy of neonatal heart and is thus
CC useful for treating cardiac insufficiency disorders. PRO1154 and PRO1186
CC stimulate adrenal cortical capillary endothelial growth, and PRO126,
CC PRO943, PRO828, PRO1068 or PRO535, PRO826, PRO819, PRO1126,
CC PRO1360 and PRO1387 induce c-fos in endothelial cells, and are thus
CC useful for treating conditions or disorders where angiogenesis would be
CC beneficial, e.g. wound healing and antagonist of this polypeptide are
CC useful for treating cancerous tumours. PRO812 inhibits vascular
CC endothelial growth factor (VEGF) stimulated proliferation of endothelial
CC cells and is thus useful for inhibiting endothelial cell growth in
CC mammals which would be beneficial in inhibiting tumour growth. PRO826,
CC PRO1068, PRO1184, PRO1346 and PRO1375 stimulate proliferation of
CC stimulated T-lymphocytes and are therapeutically useful for enhancing
CC immune response. PRO828, PRO826, PRO1068 or PRO1132 enhance survival of
CC retinal neurons cells (PRO1132 is also enhances survival/proliferation of
CC rod photoreceptor cells) and therefore are useful for treating retinal
CC disorders of injuries, e.g. retinitis pigmentosa, AMD, PRO819, PRO813
CC and PRO1066 induce proliferation of mammalian kidney mesangial cells,
CC and therefore are useful for treating kidney disorders associated with
CC decreased mesangial cell function such as Berger disease or other
CC nephropathies associated with dermatitis, herpeticiformis or Crohn's
CC disease. PRO1310, PRO844, PRO1312, PRO1192 and PRO1387 induce the
CC proliferation and/or redifferentiation of chondrocytes in culture and
CC are thus useful for treating sports injuries and arthritis. This
CC is the amino acid sequence of a novel human PRO protein.
XX
XX Sequence 343 AA:
XX
XX Query Match. 40.8%; Score 341; DB 24; Length 343;
XX Best Local Similarity 100.0%; Pred No. 0;
XX Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 MAGLNGVSIALGLVLLGAAALPRGAFAEIALPRESNITVLIKGTPTLLAKPCYIV 60
XX 1 MAGLNGVSIALGLVLLGAAALPRGAFAEIALPRESNITVLIKGTPTLLAKPCYIV 60
XX 61 SKRHITMLSTKSGERIVFTFSCSPENHPIETIQKIDCMGSGCPGEVQLPSTSLPT 120
XX 61 SKRHITMLSTKSGERIVFTFSCSPENHPIETIQKIDCMGSGCPGEVQLPSTSLPT 120
XX 61 SKRHITMLSTKSGERIVFTFSCSPENHPIETIQKIDCMGSGCPGEVQLPSTSLPT 120
XX 121 LNRFTIWDVKAHSIGLELOFSIPRLROIGPESCDGVTWHSIGIDNIVRIGFGCN 180
XX 121 LNRFTIWDVKAHSIGLELOFSIPRLROIGPESCDGVTWHSIGIDNIVRIGFGCN 180
XX 121 LNRFTIWDVKAHSIGLELOFSIPRLROIGPESCDGVTWHSIGIDNIVRIGFGCN 180
XX 181 GTVSRIKMGEGVMALHLPMFHPRNVSGSIANRSSIKRLCTIESVFESEGSATLMSANY 240
XX 181 GTVSRIKMGEGVMALHLPMFHPRNVSGSIANRSSIKRLCTIESVFESEGSATLMSANY 240
XX 181 GTVSRIKMGEGVMALHLPMFHPRNVSGSIANRSSIKRLCTIESVFESEGSATLMSANY 240
XX 241 PEGFPEDELTWQFVPAHLRAVSFLNRLSNCKEKREVERVYIGSTTNPVEVFLDEK 300
XX 241 PEGFPEDELTWQFVPAHLRAVSFLNRLSNCKEKREVERVYIGSTTNPVEVFLDEK 300
XX 241 PEGFPEDELTWQFVPAHLRAVSFLNRLSNCKEKREVERVYIGSTTNPVEVFLDEK 300
XX 301 QPNNMAGNPNLISLOGDDDAOSGRIILROFVVOHPONES 341
XX 301 QPNNMAGNPNLISLOGDDDAOSGRIILROFVVOHPONES 341
XX 301 QPNNMAGNPNLISLOGDDDAOSGRIILROFVVOHPONES 341

RESULT 10
AAV91456
ID AAV91456 standard; Protein; 709 AA.
XX
XX AAV91456;
XX
XX 29-JUN-2000 (first entry)
XX
XX Human secreted protein sequence encoded by gene 6 SEQ ID NO:129.
XX
XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
XX antiHIV; antiinflammatory; nootropic; neuroprotective; antiallergic;
XX osteopathic; antirheumatic; antibacterial; antidiabetic; antiashma;
XX antipsoriatic; cardiant; gene therapy; cancer; neurological disorder;
XX immune disease; inflammation; blood disorder; tumour.
XX
XX Homo sapiens.
XX
XX WO200006698-A1.
XX
XX 10-FEB-2000.
XX
XX 29-JUL-1999; 99WO-US17130.
XX
XX 30-JUL-1998; 98US-0094657.
XX 05-AUG-1998; 98US-0095486.
XX 06-AUG-1998; 98US-0095454.
XX 06-AUG-1998; 98US-0095455.
XX 12-AUG-1998; 98US-0096319.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Komatsu S, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;
XX Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;
XX Soppet DR, Endres GA, Emner R, Olsen HS, Mucenski M;
XX WPI: 2000-195282/17.
XX N-PSDB; AAA26351.
XX
XX New isolated human genes and the secreted polypeptides they encode,
XX useful for diagnosis and treatment of e.g. cancer, neurological
XX disorders, immune diseases, inflammation or blood disorders -
XX
XX Claim 11: Page 456-459; 634pp; English.
XX
XX The polynucleotide sequences given in AAA26346 to AAA26458 encode the
XX human secreted proteins given in AAV91451 to AAV91461. The human secreted
XX proteins can have activities based on the tissues and cells they are
XX expressed in. Examples of the activities are: cytostatic;
XX immunosuppressive; antiHIV; antiinflammatory; nootropic; neuroprotective;
XX antiallergic; osteopathic; antirheumatic; antibacterial; antidiabetic;
XX antiashma; antipsoriatic; and cardiant. The polynucleotides and their
XX corresponding secreted proteins are useful for preventing, treating or
XX ameliorating medical conditions, e.g. by protein or gene therapy. Also
XX pathological conditions can be diagnosed by determining the amount of the
XX proteins in a sample or by determining the presence of mutations in the
XX polynucleotides. Specific uses are described for each of the
XX polynucleotides, based on which tissues they are most highly expressed
XX in, and include developing products for the diagnosis or treatment of
XX cancer, tumours, neurodegenerative disorders, developmental abnormalities
XX and foetal deficiencies, blood disorders, diseases of the immune system,
XX autoimmune diseases, hepatic and renal disease, inflammation,
XX allergies, Alzheimer's and behavioural disorders, schizophrenia,
XX osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
XX transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
XX cardiovascular disorders, reproductive disorders, gastrointestinal
XX disorders, respiratory disorders and metabolic disorders. The proteins
XX or polynucleotides can also be used as food additives or preervatives.
XX The proteins are also useful for identifying their binding partners.
XX AAA26337 to AAA26345 and AAV91450 are sequences used in the
XX exemplification of the present invention.

```

XX SQ Sequence 709 AA:
Query Match 25.2%; Score 211; DB 21; Length 709;
Best Local Similarity 99.5%; Pred. No. 2,9e-197;
Matches 411; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 MAGLNCGVSYALLGVLLGAARLPGAAEAFIALPRESNITVLIKLGPPTLLAKPCYIV 60
DB 1 MAGLNCGVSYALLGVLLGAARLPGAAEAFIALPRESNITVLIKLGPPTLLAKPCYIV 60
OY 61 SKRHITMLSIKSGERIVFTFSCQSPBNHFVITQKIDCMSPCPGGEVQLPSTSLPT 120
DB 61 SKRHITMLSIKSGERIVFTFSCQSPBNHFVITQKIDCMSPCPGGEVQLPSTSLPT 120
OY 121 LNRITIMDKAKHSIGLELOFSIPRLROI GPESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRITIMDKAKHSIGLELOFSIPRLROI GPESCPDGVTHSISGRIDATVVRIGTFCSN 180
OY 181 GTVSRIKMOEGVKMLHLPMPHPRNVSGFSIANRSSIKLCITIEVFEGEGSATLMSANY 240
DB 181 GTVSRIKMOEGVKMLHLPMPHPRNVSGFSIANRSSIKLCITIEVFEGEGSATLMSANY 240
OY 241 PEGFPEDELMTWQFVVPAPHLRASVSFLNPNLSNCERKEERVEYIIPGSTTNDPFLBDK 300
DB 241 PEGFPEDELMTWQFVVPAPHLRASVSFLNPNLSNCERKEERVEYIIPGSTTNDPFLBDK 300
OY 301 OPGNAGNPNLSLQCCDDOAGSPGILRLQFOVLVQHPQESKIKIYVVDLSNRRAMSLTTE 360
DB 301 OPGNAGNPNLSLQCCDDOAGSPGILRLQFOVLVQHPQESKIKIYVVDLSNRRAMSLTTE 360
OY 361 PRPVQSRKRVPGCFVCLSESRFCSSNLTLLTSGSKHISFLCDLTRLMMNVK 413
DB 361 PRPVQSRKRVPGCFVCLSESRFCSSNLTLLTSGSKHISFLCDLTRLMMNVK 413

RESULT 11
ID AAY91593 standard; Protein; 159 AA.
AC AAY91593;
XX
XX 29-JUN-2000 (first entry)
XX
DE Human secreted protein sequence encoded by gene 6 SEQ ID NO:266.
XX
XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
XX antiHIV; antiinflammatory; nootropic; neuroprotective; anti allergy;
XX osteopathic; antiarthritic; antibacterial; antidiabetic; antiashma;
XX antiparasitic; cardiant; gene therapy; cancer; neurological disorder;
XX immune disease; inflammation; blood disorder; tumour.
XX
OS Homo sapiens.
XX
XX WO200006698-A1.
XX
XX 10-FEB-2000.
XX
XX 29-JUL-1999; 99WO-US17130.
XX
XX 30-JUL-1998; 98US-0094657.
XX 05-AUG-1998; 98US-0095486.
XX 06-AUG-1998; 98US-0095454.
XX 06-AUG-1998; 98US-0095455.
XX 12-AUG-1998; 98US-0096319.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;
XX Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;
XX Soppet DR, Endreess GA, Ebner R, Olsen HS, Mucenski M;
XX
XX WPI; 2000-195282/17.

```

```

XX New isolated human genes and the secreted polypeptides they encode,
XX useful for diagnosis and treatment of e.g. cancers, neurological
XX disorders, immune diseases, inflammation or blood disorders -
XX
XX Disclosure: Page 18; 634pp; English.
XX
XX The polynucleotide sequences given in AAA26346 to AAA26458 encode the
XX human secreted proteins given in AAY91451 to AAY91691. The human secreted
XX proteins can have activities based on the tissues and cells they are
XX expressed in. Examples of the activities are: cytostatic;
XX immunosuppressive; antiHIV; antiinflammatory; nootropic; neuroprotective;
XX antiallergic; osteopathic; antiarthritic; antibacterial; antidiabetic;
XX antiashma; antiparasitic; and cardiant. The polynucleotides and their
XX corresponding secreted proteins are useful for preventing, treating or
XX ameliorating medical conditions, e.g. by protein or gene therapy. Also
XX pathological conditions can be diagnosed by determining the amount of the
XX proteins in a sample or by determining the presence of mutations in the
XX polynucleotides. Specific uses are described for each of the
XX polynucleotides, based on which tissues they are most highly expressed
XX in, and include developing products for the diagnosis or treatment of
XX cancer, tumours, neurodegenerative disorders, developmental abnormalities
XX and foetal deficiencies, blood disorders, diseases of the immune system,
XX autoimmune diseases, hepatic and renal disease, inflammation,
XX allergies, Alzheimer's and behavioural disorders, schizophrenia,
XX osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
XX transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
XX cardiovascular disorders, reproductive disorders, gastrointestinal
XX disorders, respiratory disorders and metabolic disorders. The proteins
XX or polynucleotides can also be used as food additives or preservatives.
XX The proteins are also useful for identifying their binding partners.
XX AAA26337 to AAA26345 and AAY91450 are sequences used in the
XX exemplification of the present invention.
XX
XX Sequence 159 AA:
XX
XX Query Match 19.0%; Score 159; DB 21; Length 159;
XX Best Local Similarity 100.0%; Pred. No. 6,3e-147;
XX Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 30 FEIALPRESNITVLIKLGPPTLLAKPCYIVISKRHITMLSIKSGRIIVFTFSCQSPBNHF 89
DB 1 FEIALPRESNITVLIKLGPPTLLAKPCYIVISKRHITMLSIKSGRIIVFTFSCQSPBNHF 60
OY 90 VIEIQKNIDCMSPCPGGEVQLPSTSLPTLNRFIMDVAKHSIGLELOFSIPRLROI 149
DB 61 VIEIQKNIDCMSPCPGGEVQLPSTSLPTLNRFIMDVAKHSIGLELOFSIPRLROI 120
OY 150 PGESCPDGVTHSISGRIDATVVRIGTFCSSNGTVSRIM 188
DB 121 PGESCPDGVTHSISGRIDATVVRIGTFCSSNGTVSRIM 159

RESULT 12
ID ABG00296 standard; Protein; 143 AA.
XX
XX ABG00296;
XX
XX 13-FEB-2002 (first entry)
XX
XX Novel human diagnostic protein #287.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.

```


XX 17-JAN-2001; 2001WO-US01354.
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0225729.
 PR 22-AUG-2000; 2000US-0226811.
 PR 22-AUG-2000; 2000US-0226868.
 PR 23-AUG-2000; 2000US-0227182.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0228287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0232081.
 PR 14-SEP-2000; 2000US-0233397.
 PR 14-SEP-2000; 2000US-0233398.
 PR 14-SEP-2000; 2000US-0233399.
 PR 14-SEP-2000; 2000US-0233400.
 PR 14-SEP-2000; 2000US-0233401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234977.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.

PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241826.
 PR 08-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0250300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251988.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM,
 WPI; 2001-483426/52.
 N-PSDB; AAK63123.

XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
XX Claim 11; SEQ ID NO 17935; 3071bp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
XX Sequence 116 AA;
SQ
Query Match 8.9%; Score 74; DB 22; Length 116;
Best Local Similarity 100.0%; Pred. No. 7-8e-64;
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 343 KIYVVDLSNERAMSLTIEPRPVKQSRKRVGCFVCLIESRTCSNLTLTGSKHKISFLCD 402
DB 14 KIYVVDLSNERAMSLTIEPRPVKQSRKRVGCFVCLIESRTCSNLTLTGSKHKISFLCD 73
QY 403 ||||| 416
DB 74 DLTRLMNMNVEKITS 87
RESULT 15
AAU69483
ID AAU69483 standard; Protein; 167 AA.
XX
XX AAU69483;
AC
XX
XX 29-JAN-2002 (first entry)
DT
XX
XX Human purified secretory polypeptide #52.
DE
XX
XX Human; purified secretory polypeptide; cell proliferative disorder;
KW cancer; immune system disorder; neurological disorder; mental disorder;
KW motor neuron disorder; demyelinating disorder; neuromuscular disorder;
KW central nervous system disorder; enzyme linked immunosorbent assay;
KW ELISA; gene therapy.
XX
XX
XX Homo sapiens.
OS
XX
XX MO200162918-A2.
PN
XX
XX 30-AUG-2001.
PD
XX
XX 01-FEB-2001; 2001WO-US03465.
PF
XX
XX 24-FEB-2000; 2000US-185215P.
PR 24-FEB-2000; 2000US-185216P.
PR 16-MAY-2000; 2000US-205232P.
PR 17-MAY-2000; 2000US-205286P.
PR 17-MAY-2000; 2000US-205287P.
PR 17-MAY-2000; 2000US-205323P.
PR 17-MAY-2000; 2000US-205324P.
XX
XX (INCY-) INCYTE GENOMICS INC.
PA

XX Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC;
PI Chen A, D'sa SA, Amshey S, Dahl CR, Dam TC, Daniels SE;
PI Dufour GE, Flores V, Fong WT, Greenawalt LB, Hillman JL, Jones AL;
PI Liu TF, Roseberry AM, Rosen BH, Russo PD, Stockdreher TK, Daffo A;
PI Wright RJ, Yap PE, Yu JY, Bradley DL, Bratcher SR, Chen W;
PI Cohen HJ, Hodgson DM, Lincoln SE;
XX
XX WPI; 2001-648217/74.
DR
XX
XX Nucleic acids encoding secretory polypeptides, useful in genetic
PT diagnosis and therapy -
XX
XX
XX Disclosure; Page 221; 237pp; English.
XX
XX Sequences AAU69432-AAU69511 represent purified secretory polypeptides of
CC the invention. The polypeptides and their associated polynucleotides can
CC be used in the treatment, prevention and diagnosis of diseases associated
CC with inappropriate secretory protein expression. These diseases include
CC cell proliferative disorders such as atherosclerosis and psoriasis,
CC cancers such as leukaemia and melanoma, immune system disorders such as
CC asthma and diabetes mellitus, neurological disorders such as epilepsy and
CC Parkinson's disease, mental disorders such as schizophrenia and seasonal
CC affective disorder (SAD), motor neuron disorders such as amyotrophic
CC lateral sclerosis, demyelinating disorders such as multiple sclerosis,
CC central nervous system disorders such as mental retardation and
CC neurofibromatosis and neuromuscular disorders such as cerebral palsy and
CC muscular dystrophy. Target polynucleotides in a sample can be detected by
CC hybridising the sample with a probe sequence complementary to the target
CC polynucleotide, under conditions in which a hybridisation complex is
CC formed, and detecting the presence or absence of the complex. The
CC polypeptides may also be used as antigens in the production of antibodies
CC against secretory proteins and in assays to identify modulators of
CC protein expression and activity. The antibodies may also be used as
CC diagnostic agents for detecting the presence of the sequences of the
CC invention in samples e.g. by enzyme linked immunosorbent assay (ELISA).
XX
XX Sequence 167 AA;
SQ
Query Match 4.7%; Score 39; DB 22; Length 167;
Best Local Similarity 100.0%; Pred. No. 2e-29;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 107 GEVQLQPSSTSLPTLNRTFTIMDVKAHKSIGLEQPSIR 145
DB 23 GEVQLQPSSTSLPTLNRTFTIMDVKAHKSIGLEQPSIR 61

Search completed: February 20, 2004, 19:46:32
Job time : 89 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2004, 19:11:57 ; Search time 47 Seconds
(without alignments)
1710.575 Million cell updates/sec

Title: US-09-899-569a-4

Perfect score: 836
Sequence: 1 MAGLNCGVSIALLGVLLGA.....SSKQTDIPLNTQEPMEBAE 836

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	1.1	293	2	H70713
2	8	1.0	144	2	TS0500
3	8	1.0	187	1	Q0EC30
4	8	1.0	187	2	B98089
5	8	1.0	187	2	D85934
6	8	1.0	203	2	T20750
7	8	1.0	213	2	B71607
8	8	1.0	232	2	AD2542
9	8	1.0	262	2	B83827
10	8	1.0	390	2	T41199
11	8	1.0	404	2	A35407
12	8	1.0	550	1	VGBE18
13	8	1.0	603	2	S70849
14	8	1.0	682	2	F83228
15	8	1.0	74	2	T18489
16	7	0.8	74	2	T47434
17	7	0.8	83	2	AE0713
18	7	0.8	101	2	AB3548
19	7	0.8	104	2	A33731
20	7	0.8	107	2	C83265
21	7	0.8	107	2	C86477
22	7	0.8	119	2	S08077
23	7	0.8	129	2	B90105
24	7	0.8	129	2	AC2036
25	7	0.8	132	2	S63978
26	7	0.8	132	2	B83299
27	7	0.8	140	1	F2MU10
28	7	0.8	140	2	G83318
29	7	0.8	149	2	C72611

30	7	0.8	152	2	F64441	molybdenum cofacto
31	7	0.8	154	1	UQDOR7	ubiquitin / riboso
32	7	0.8	156	2	B84991	30S ribosomal prot
33	7	0.8	162	2	T17047	NADH2 dehydrogenas
34	7	0.8	162	2	T17045	NADH2 dehydrogenas
35	7	0.8	163	2	T42696	hypothetical prote
36	7	0.8	168	2	PH0139	M protein-like mol
37	7	0.8	168	2	D71266	hypothetical prote
38	7	0.8	171	2	T41924	hypothetical prote
39	7	0.8	174	2	T15117	NADH2 dehydrogenas
40	7	0.8	176	2	D97084	enzyme of dihydrof
41	7	0.8	186	2	A75281	hypothetical prote
42	7	0.8	187	2	AF3009	hypothetical prote
43	7	0.8	188	2	JC5154	K-ras protein - tu
44	7	0.8	189	2	JC4072	virulence-associat
45	7	0.8	195	2	H98274	hypothetical prote

ALIGNMENTS

RESULT 1
H70713
Hypothetical protein RV1509 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 28-Jul-2000
C:Accession: H70713
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:9829587; PMID:9634230
A:Accession: H70713
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-293 <COL>
A:Cross-references: GB:Z79701; GB:AL123456; NID:93261635; PIDN:CAB02023.1; PID:e264138; I
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV1509
C:Superfamily: Mycobacterium tuberculosis hypothetical protein RV1509

Query Match 1.1%; Score 9; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 674 VGGGVLLLS 682
DB 67 VGGGVLLLS 75

RESULT 2
TS0500
Hypothetical protein T22D6.90 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 08-Sep-2000
C:Accession: TS0500
R:Bevan, M.; Terry, N.; Ardiles, W.; Buysaert, C.; Dasseville, R.; De Clerck, R.; De
ewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25101
A:Accession: TS0500
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-144 <BEV>
A:Cross-references: EMBL:AL357612
A:Experimental source: cultivar Columbia; BAC clone T22D6
C:Genetics:
A:Map position: 5
A:Note: T22D6.90
C:Superfamily: Arabidopsis thaliana hypothetical protein T22D6.90

Query Match 1.0%; Score 8; DB 2; Length 144;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 692 KKKKKKTN 699
Db 81 KKKKKKTN 88

RESULT 3

prepilin peptidase dependent protein B precursor - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C:Date: 31-Mar-1988 #sequence_revision 05-Dec-1997 #text_change 01-Mar-2002

C:Accession: B65065; C24137

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; PMID:97426617; PMID:9278503

A:Accession: B65065

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-187 <BLAT>

A:Cross-references: GB:AE000366; GB:U00096; NID:g1789185; PIDN:AACT5864.1; PID:g1789189;

A:Experimental source: strain K-12, substrain MGL655

R:Finch, P.W.; Wilson, R.E.; Brown, K.; Hickson, I.D.; Tomkinson, A.E.; Emerson, P.T.

Nucleic Acids Res. 14, 4437-4451, 1986

A:Title: Complete nucleotide sequence of the Escherichia coli recC gene and of the thyA-

A:Reference number: A93625; PMID:86232583; PMID:3520484

A:Accession: C24137

A:Molecule type: DNA

A:Residues: 'MVPCFRCQYLSMPTIGCATLARSRRERPTFACFLMPAQYRMGRAYSLOKINGRVAGGFAVGTPLA

A:Cross-references: GB:X03966; NID:g42684; PIDN:CAA27601.1; PID:g42686

C:Genetics:

A:Gene: ppdB

A:Map position: 61 min

C:Superfamily: prepilin peptidase dependent protein B precursor

Query Match 1.0%; Score 8; DB 1; Length 187;

Best Local Similarity 100.0%; Pred. No. 9;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLIGAAR 22

Db 22 VLLIGAAR 29

RESULT 4

prepilin peptidase dependent protein B [imported] - Escherichia coli (strain O157:H7, su

C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C:Accession: B98089

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasaawara, N.; Yasunaga, T.; Kohara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno

A:Reference number: A99629; PMID:21156231; PMID:11258796

A:Accession: B98089

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-187 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA037105.1; PID:g13363154; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RIMD 0509952

C:Genetics:

Qy 15 VLLIGAAR 22

Db 22 VLLIGAAR 29

RESULT 5

prepilin peptidase dependent protein B [imported] - Escherichia coli (strain O157:H7, su

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

C:Accession: D85934

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,

Iller, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; Dialianta, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; PMID:21074935; PMID:11206551

A:Accession: D85934

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-187 <STO>

A:Cross-references: GB:AE005174; NID:g12517310; PIDN:AG57936.1; GSPDB:GN00145; UWGP:Z414

A:Experimental source: strain O157:H7, substrain EDJ933

C:Genetics:

A:Gene: ppdB

C:Superfamily: prepilin peptidase dependent protein B precursor

Query Match 1.0%; Score 8; DB 2; Length 187;

Best Local Similarity 100.0%; Pred. No. 9;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLIGAAR 22

Db 22 VLLIGAAR 29

RESULT 6

hypothetical protein F11A5.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-Jan-2000

C:Accession: T20750

R:Gardner, A.

submitted to the EMBL Data Library, March 1997

A:Reference number: Z19319

A:Accession: T20750

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-203 <WIL>

A:Cross-references: EMBL:Z92830; PIDN:CAB07357.1; GSPDB:GN00023; CESP:F11A5.4

A:Experimental source: clone F11A5

C:Genetics:

A:Gene: CESP:F11A5.4

A:Map position: 5

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

Query Match 1.0%; Score 8; DB 2; Length 203;

Best Local Similarity 100.0%; Pred. No. 9.7;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 690 CVYKXXXX 697

Db 180 CVYKXXXX 187

RESULT 7

hypothetical protein PFB0740C - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000

C:Accession: B71607

R:Gardner, M.J.; Jectel, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;

Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.;

Science 282, 1126-1132, 1998
 A>Title: Chromosome 2 sequence of the human malaria parasite *Plasmodium falciparum*.
 A:Reference number: A71600; MUID:99021743; PMID:9804551
 A:Accession: B71607
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-213 <GAR>
 A:Cross-references: GB:AE001415; GB:AE001362; NID:g3845264; PIDN:AACT1937.1; PID:g384526
 A:Experimental source: clone 3D7
 C:Genetics:
 A:Gene: PFB0740C

Query Match 1.0%; Score 8; DB 2; Length 213;
 Best Local Similarity 100.0%; Pred.No. 10;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 692 KKKKKKTN 699
 DB 5 KKKKKKTN 12

RESULT 8
 AD2542
 Hypothetical protein all7623 (imported) - *Nostoc* sp. (strain PCC 7120) plasmid pCC7120be
 C:Species: *Nostoc* sp. PCC 7120
 A>Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
 C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 12-May-2003
 C:Accession: AD2542
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Saemoto, S.; Matsumoto, A.; Itiguchi, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena*
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AD2542
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-232 <KUR>
 A:Cross-references: GB:AP003602; PIDN:BA877266.1; PID:g17134708; GSPDB:GN00181
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: all7623
 A:Genome: plasmid
 C:Superfamily: bacitracin transport permease

Query Match 1.0%; Score 8; DB 2; Length 232;
 Best Local Similarity 100.0%; Pred.No. 11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 666 LTVILIAA 673
 DB 179 LTVILIAA 166

RESULT 9
 B83827
 Hypothetical protein BH1418 (imported) - *Bacillus halodurans* (strain C-125)
 C:Species: *Bacillus halodurans*
 C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C:Accession: B83827
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira, N.
 Nucleic Acids Res. 28, 4317-4331, 2000
 A>Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
 A:Reference number: AB3650; MUID:20512582; PMID:11058132
 A:Accession: B83827
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-262 <STO>
 A:Cross-references: GB:AF001512; GB:BA000004; NID:g10174030; PIDN:BA805137.1; GSPDB:GN00
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BH1418
 C:Superfamily: hypothetical protein MJ0933

Query Match 1.0%; Score 8; DB 2; Length 262;
 Best Local Similarity 100.0%; Pred.No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ALLGVLL 18
 DB 5 ALLGVLL 12

RESULT 10
 T41199
 dom34 protein homolog - fission yeast (*Schizosaccharomyces pombe*)
 C:Species: *Schizosaccharomyces pombe*
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 C:Accession: T41199
 R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Jimenez Martinez, J.
 submitted to the EMBL Data Library, July 1999
 A:Reference number: Z21978
 A:Accession: T41199
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-390 <MOO>
 A:Cross-references: EMBL:AL109736; PIDN:CA852153.1; GSPDB:GN00068
 A:Experimental source: strain 972h-; cosmid c1885
 C:Genetics:
 A:Gene: SPBC1885.06
 A:Map position: 3
 A:introns: 27/2; 45/2; 145/1; 180/3
 C:Superfamily: cell division protein MJ0174

Query Match 1.0%; Score 8; DB 2; Length 390;
 Best Local Similarity 100.0%; Pred.No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 647 SGKQDL 654
 DB 352 SGKQDL 359

RESULT 11
 A35407
 tryptophan synthase (EC 4.2.1.20) beta chain - *Thermus aquaticus*
 C:Species: *Thermus aquaticus*
 C>Date: 14-Sep-1990 #sequence_revision 14-Sep-1990 #text_change 16-Jul-1999
 C:Accession: A35407
 R:Koyama, Y.; Furukawa, K.
 J. Bacteriol. 172, 3490-3495, 1990
 A>Title: Cloning and sequence analysis of tryptophan synthetase genes of an extreme therm
 to competent *T. thermophilus* cells.
 A:Reference number: A35407; MUID:90264352; PMID:2188962
 A:Accession: A35407
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-404 <KOY>
 A:Cross-references: GB:M32108; NID:g155132; PIDN:AAA27508.1; PID:g155133
 A:Experimental source: strain HB27
 C:Genetics:
 A:Gene: trpB
 C:Function:
 A:Description: catalyzes conversion of indoleglycerol phosphate and serine to tryptophan
 A:Pathway: tryptophan biosynthesis
 A>Note: cofactor pyridoxal phosphate; last step in pathway
 C:Superfamily: tryptophan synthase beta chain; tryptophan synthase beta chain homology
 C:Keywords: carbon-oxygen lyase; hydro-lyase; phosphoprotein; pyridoxal phosphate; trypt
 F;12-397/Domain: tryptophan synthase beta chain homology <TRPB>
 F;94/Active site: His #status predicted
 F;95/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Query Match 1.0%; Score 8; DB 2; Length 404;
 Best Local Similarity 100.0%; Pred.No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 670 LIAAVGG 677

Db 235 LIAAVGGG 242

RESULT 12

VGBE18

glycoprotein E - human herpesvirus 1

N:Alternate names: US8

C:Species: human herpesvirus 1

C>Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 16-Jul-1999

C:Accession: A03733; A45696

R:McGeoch, D.J.; Dolan, A.; Donald, S.; Rixon, F.J.

J. Mol. Biol. 181, 1-13, 1985

A>Title: Sequence determination and genetic content of the short unique region in the ge

A:Reference number: A00656; MUID:85160822; PMID:2984429

A:Accession: A03733

A:Molecule type: DNA

A:Residues: 1-550 <MCG>

A:Cross-references: GB:X02138, NID:959865; PIDN:CAA26062.1; PID:959882

A:Experimental source: strain 17

R:Georgopoulou, U.; Michaelidou, A.; Roizman, B.; Mavromara-Nazos, P.

J. Virol. 67, 3961-3968, 1993

A>Title: Identification of a new transcriptional unit that yields a gene product within

A:Reference number: A45696; MUID:93287213; PMID:8389914

A:Accession: A45696

A:Molecule type: DNA

A:Status: preliminary

A:Cross-references: GB:S62895; NID:9386127; PIDN:AA27080.1; PID:9386128

A:Experimental source: R35

A>Note: sequence extracted from NCBI backbone (NCBIN:133646, NCBIPI:133647)

C:Superfamily: herpesvirus glycoprotein E

C:Keywords: glycoprotein

F:124,243,501/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 1.0%; Score 8; DB 1; Length 550;

Best Local Similarity 100.0%; Pred. No. 23;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 679 LLLSALGL 686
Db 428 LLLSALGL 435

RESULT 13

S70849

cholinesterase (EC 3.1.1.8) - mouse

N:Alternate names: butyrylcholine esterase

C:Species: Mus musculus (house mouse)

C>Date: 28-Oct-1996 #sequence_revision 08-Nov-1996 #text_change 18-Jun-1999

C:Accession: S70849; S15680; A39768

R:Taylor, P.

submitted to the EMBL Data Library, August 1992

A:Reference number: S70849

A:Accession: S70849

A:Molecule type: nucleic acid

A:Residues: 1-603 <TAY>

A:Cross-references: EMBL:M99492; NID:9191579; PIDN:AAA37328.1; PID:9191580

R:Rachinsky, T.L.; Camp, S.; Li, Y.; Ekstrom, T.J.; Newton, M.; Taylor, P.

Neuron 5, 317-327, 1990

A>Title: Molecular cloning of mouse acetylcholinesterase: tissue distribution of alterna

A:Reference number: JH0314; MUID:90380429; PMID:2400605

A:Accession: S15680

A:Status: nucleic acid sequence not shown

A:Molecule type: nucleic acid

A:Residues: 30-128, P, 130-603 <RAC>

A:Cross-references: EMBL:M99492

R:Arpagaus, M.; Chatonnet, A.; Masson, P.; Newton, M.; Vaughan, T.A.; Bartels, C.F.; Nog

J. Biol. Chem. 266, 6966-6974, 1991

A>Title: Use of the polymetase chain reaction for homology probing of butyrylcholinester

A:Reference number: A39768; MUID:91201348; PMID:2016308

A:Accession: A39768

A:Status: preliminary

A:Molecule type: DNA
A:Residues: 97-128, P, 130-237 <ARP>
A:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase; glycoprotein
F:57-557/Domain: cholinesterase homology <CHE>

Query Match 1.0%; Score 8; DB 2; Length 603;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 613 TRAEIRFS 620
Db 486 TRAEIRFS 493

RESULT 14

hypothetical protein PA3340 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: F83228

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bri

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapidig, K.; Llm,

.; Loly, S.; Olson, M.V.

Nature 406, 959-964, 2000

A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathog

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: F83228

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-682 <STO>

A:Cross-references: GB:AE004756; GB:AE004091; NID:99949466; PIDN:AA606728.1; GSPDB:GN0013

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA3340

Query Match 1.0%; Score 8; DB 2; Length 682;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ALLGVLL 18
Db 364 ALLGVLL 371

RESULT 15

T18489

hypothetical protein C0820w - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000

C:Accession: T18489

R:Lawson, D.; Bowman, S.; Barrett, B.

submitted to the EMBL Data Library, August 1997

A:Reference number: Z18935

A:Accession: T18489

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-4981 <LAW>

A:Cross-references: EMBL:Z98551; NID:e1331903; PID:e1331910; PIDN:CA11128.1

C:Genetics:

A:Map position: 3

A>Note: C0820w

Query Match 1.0%; Score 8; DB 2; Length 4981;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 692 KKKKKKTN 699
Db 1439 KKKKKKTN 1446

Search completed: February 20, 2004, 19:50:19

Mon Feb 23 08:34:08 2004

us-09-899-569a-4.011.rpr

Job time : 49 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2004, 13:55:12 ; Search time 27 Seconds

(without alignments)
1456.086 Million cell updates/sec

Title: US-09-899-569a-4

Perfect score: 836
Sequence: 1 MAGLNCGVSLALGVLLGA.....SSKQDIPLNTQBPMPAP 836

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size: 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	1.1	173	PSAL_MASLA	O31126 mastigoclad
2	9	1.1	293	YF09_MYCTU	P71788 mycobacteri
3	8	1.0	187	PPDB_ECCLI	P08371 escherichia
4	8	1.0	404	TRPB_THETH	P16609 thermus the
5	8	1.0	485	ZDH1_HUMAN	O8wtx9 homo sapien
6	8	1.0	485	ZDH1_MOUSE	O8wcn9 mus musculu
7	8	1.0	550	USHA_SALPU	O9fn37 salmoneilla
8	8	1.0	550	VGLE_HSV1	P04488 herpes simp
9	8	1.0	603	CHLE_MOUSE	O03311 mus musculu
10	8	1.0	78	R27A_DICDI	P14797 dictyosteli
11	7	0.8	104	SLIB_RAT	P09916 rattus norv
12	7	0.8	132	SMDA_STRCS	P56406 streptomyce
13	7	0.8	140	PSBR_ARATH	P27202 arabidopsis
14	7	0.8	141	PSBR_BRACH	P49108 brassica ca
15	7	0.8	152	MOAC_METVA	O58535 methanococ
16	7	0.8	152	PSAL_METVA	O74699 quillardi
17	7	0.8	156	RS7_BUCAL	P57594 buchnera ap
18	7	0.8	156	RS7_BUCAL	P57594 buchnera ap
19	7	0.8	174	NU6M_PAPHA	O9xxxs paglo hamad
20	7	0.8	188	RASK_MELDA	P78800 meleagris g
21	7	0.8	216	SSPN_MOUSE	O62147 mus musculu
22	7	0.8	219	ATPD_CHLRE	O42687 chlamydomon
23	7	0.8	219	YIIL_STRCO	O86576 streptomyce
24	7	0.8	227	COX2_SITGR	P28879 sitophilus
25	7	0.8	243	SSPN_HUMAN	O14714 homo sapien
26	7	0.8	265	BXB2_HAETN	P19391 haemophilus
27	7	0.8	265	BXB2_HAETN	P22235 haemophilus
28	7	0.8	278	RCEL_CHRYI	P51762 chromatium
29	7	0.8	281	YE94_XYLFA	O9nd84 xyloella fas
30	7	0.8	285	IF2B_YEAST	P08064 saccharomyc
31	7	0.8	300	ERA_MYCLB	O49768 mycobacteri
32	7	0.8	312	IF2B_DROME	P41375 drosophila
33	7	0.8	328	P111_HUMAN	O9dbk5 homo sapien

34	7	0.8	331	IF2B_MOUSE	O99145 mus musculu
35	7	0.8	332	P111_MOUSE	O9cex5 mus musculu
36	7	0.8	333	IF2B_HUMAN	P20042 homo sapien
37	7	0.8	333	IF2B_RABIT	P41035 oryctolagus
38	7	0.8	335	BTUC_YERPE	O8zdx4 yerinia pe
39	7	0.8	335	LEP4_MYXXA	O30387 myxococcus
40	7	0.8	343	YD57_METVA	O58752 methanococ
41	7	0.8	346	YOC2_CABEL	O09261 caenorhabdi
42	7	0.8	371	RLX1_SALTY	P14492 salmoneilla
43	7	0.8	380	CYB_THOIS	O9xw9 thomascomye
44	7	0.8	381	CYB_AULMI	O9xw4 auliscomys
45	7	0.8	381	CYB_ELIMO	O9xw3 ellimodontic

ALIGNMENTS

RESULT 1
PSAL_MASLA STANDARD; PRT; 173 AA.
ID PSAL_MASLA
AC O31126;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Photosystem I reaction center subunit XI (PSI-X) (PSI subunit V).
GN PSAL.
OS Mastigocladus laminosus (Fischerella sp.).
OC Bacteria; Cyanobacteria; Stigonematales; Mastigocladus.
OX NCBI_TaxID=83541;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=PCC.7605;
RA He Z.-Y., Chitnis P.R., Nechushtal R.;
RT "Molecular cloning of the psal gene for photosystem I subunit XI from
the thermophilic cyanobacterium Mastigocladus laminosus.";
RL (In) Plant Gene Register PGR98-025.
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. CELLULAR
THYLAKOID MEMBRANE (PROBABLE).
CC -I- SIMILARITY: BELONGS TO THE PSAL FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
DR EMBL: AF030003; AAC04841.1; -
DR HAMAP: MF_00447; -; 1.
DR InterPro: IPR003757; PSI_Psal.
DR Pfam: PF02605; Psal; 1.
DR ProDom: PD005947; PSI_Psal; 1.
KW Photosystem I; Photosynthesis; Transmembrane.
FT TRANSMEM 92 112
FT TRANSMEM 142 162
SQ SEQUENCE 173 AA; 18444 MW; 24E04D10B908715D CRC64;
Query Match 1.1%; Score 9; DB 1; Length 173;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 680 LLSALGLI 688
95 LLSALGLI 103
RESULT 2
YF09_MYCTU STANDARD; PRT; 293 AA.
ID YF09_MYCTU
AC P71788;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)

```

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein RV1509.
GN RV1509 OR MT1557 OR MTCV277.31.
OS Mycobacterium tuberculosis.
OC Bacteria: Actinobacteria: Actinobacteridae: Actinomycetales:
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OK NCBI_TaxId=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Broesch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Baaham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornby T., Jorgels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sultson J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwim M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M.D., Salzberg S.L.,
RA Delcher A., Uterback T., Weidman J., Kouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z79701; CAB02023.1; -.
CC EMBL: AE007023; AAK45824.1; -.
CC DR EMBL; AE007023; AAK45824.1; -.
CC DR PIR; H70713; H70713.
CC DR TIGR; MT1557; -.
CC DR Tuberculist; RV1509; -.
CC KM Hypothetical protein; Complete proteome.
CC SQ SEQUENCE 293 AA; 33055 MW; 7393B5725FE71E7 CRC64;
Query Match 1.1%; Score 9; DB 1; Length 293;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 674 VGGGVILLLS 682
DB 67 VGGGVILLLS 75

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=86232583; PubMed=3520484;
RA Finch P.W., Wilson R.E., Brown K., Hickson I.D., Tomkinson A.E.,
RA Emerson P.T.;
RT "Complete nucleotide sequence of the Escherichia coli recC gene and
RT of the thvA-recC intergenic region."
RL Nucleic Acids Res. 14:4437-4451(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Klapatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [3]
RP IDENTIFICATION.
RX MEDLINE=95020523; PubMed=7934814;
RA Hobbs M., Matlick J.S.;
RT "Common components in the assembly of type 4 fimbriae, DNA transfer
RT systems, filamentous phage and protein secretion apparatus: a general
RL Mol. Microbiol. 10:233-243(1993).
CC -----
CC - FUNCTION: NOT YET KNOWN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X03966; CAA27601.1; ALT_INIT.
CC EMBL: U29581; AAB40472.1; -.
CC DR EMBL; AE000366; AAC75864.1; -.
CC DR PIR; B65065; Q0EC30.
CC DR Ecogen; EG11156; ppdB.
CC DR InterPro; IPR001120; Prok N methyl.
CC DR PROSITE; PS00409; PROKAR_NTER_METHYL. 1.
CC KM Methylation; Complete proteome.
CC FT PROPEP 1 7 BY SIMILARITY.
CC FT CHAIN 8 187 PREPELIN PEPTIDASE DEPENDENT PROTEIN B.
CC FT MOD RES 8 8 METHYLATION (BY SIMILARITY).
CC SQ SEQUENCE 187 AA; 20519 MW; 433303D846D786F0 CRC64;
Query Match 1.0%; Score 8; DB 1; Length 187;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 15 VLLGAAR 22
DB 22 VLLGAAR 29

```

```

RESULT 4
ID TRPB_THETH STANDARD; PRT; 404 AA.
AC P16609;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tryptophan synthase beta chain (EC 4.2.1.20).
GN TRPB.
OS Thermus thermophilus.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OK NCBI_TaxId=274;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HB27;

```

RA MEDLINE=90264352; PubMed=2188962;
RA Koyama Y., Furukawa K.;
RT "Cloning and sequence analysis of tryptophan synthetase genes of an
RT extreme thermophile, Thermus thermophilus H27: plasmid transfer from
RT replica-plated *Escherichia coli* recombinant colonies to competent *T.*
RT *thermophilus* cells";
RL J. Bacteriol. 172:3490-3495 (1990).
RN [2]
RP SEQUENCE OF 1-45 FROM N.A.
RC STRAIN=HB27;
RX MEDLINE=91130853; PubMed=2283046;
RA Koyama Y., Arikawa Y., Furukawa K.;
RT "A plasmid vector for an extreme thermophile, *Thermus thermophilus*,";
RL FEMS Microbiol. Lett. 60:97-101 (1990).
CC -1- FUNCTION: The beta subunit is responsible for the synthesis of L-
CC tryptophan from indole and L-serine.
CC -1- CATALYTIC ACTIVITY: L-serine + 1-(indol-3-yl)glycerol 3-phosphate
CC = L-tryptophan + glyceraldehyde 3-phosphate.
CC -1- COFACTOR: Pyridoxal phosphate (By similarity).
CC -1- PATHWAY: Tryptophan biosynthesis; fitch (last) step.
CC -1- SUBUNIT: Tetramer of two alpha and two beta chains (By
CC similarity).
CC -1- SIMILARITY: BELONGS TO THE TRPB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32108; AAA27508.1; -;
DR EMBL; X58673; CAA41527.1; -;
DR HSSP; P00933; 2MSY.
DR HAMAP; MF_00133; -; 1.
DR InterPro; IPR001926; B6_enzyme_beta.
DR InterPro; IPR006653; Trp_synth_b_rel.
DR InterPro; IPR006654; Trp_synth_beta.
DR Pfam; PF00291; PALP; 1.
DR TIGRFAMs; TIGR00263; trpb; 1.
DR PROSITE; PS00168; TRP_SYNTHASE_BETA; 1.
RM Tryptophan biosynthesis; Pyridoxal phosphate; Lyase.
FT BINDING 95 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 404 AA; 43809 MW; 74D09A1EEC2A0466 CRC64;
Query Match 1.0%; Score 8; DB 1; Length 404;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 670 LIAAVGGG 677
DB 235 LIAAVGGG 242
RESULT 5
ZDHL_HUMAN
ID_ZDHL_HUMAN STANDARD; PRT; 485 AA.
AC Q8WTX9; O15461;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger DHHC domain containing protein 1 (zinc finger protein 377)
DE (DHHC-domain-containing cysteine-rich protein 1).
GN ZDHHC1 OR ZNF377 OR C16ORF1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
CX NCB1_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=uterus;
RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Ditschenko L., Marzina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein W.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.J.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Boulford G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE OF 1-293 FROM N.A., AND TISSUE SPECIFICITY.
RC TISSUE=Pancreas;
RX MEDLINE=99321009; PubMed=10395086;
RA Puttina T., Wong P., Gentileman S.;
RT "The DHHC domain: a new highly conserved cysteine-rich motif";
RL Mol. Cell. Biochem. 195:219-226 (1999).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: Expressed at high levels in fetal lung, kidney
CC and heart. Expressed at lower levels in adult pancreas and lung.
CC -1- SIMILARITY: Contains 1 DHHC-type zinc finger.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; BC021908; AAH21908.1; -;
DR EMBL; U90653; AAB86591.2; -;
DR Gene; HGNC:17916; ZDHHC1.
DR InterPro; IPR001594; Znf_DHHC.
DR Pfam; PF01529; ZF-DHHC; 1.
DR PRODOM; PD003041; Znf_DHHC; 1.
DR PROSITE; PS0216; ZF-DHHC; 1.
RM Transmembrane; Zinc-finger.
FT TRANSMEM 53
FT TRANSMEM 78
FT TRANSMEM 78
FT TRANSMEM 186
FT TRANSMEM 242
FT ZN_FING 134
FT ZN_FING 184
SQ SEQUENCE 485 AA; 54818 MW; 6B75B07D7D82F358 CRC64;
Query Match 1.0%; Score 8; DB 1; Length 485;
Best Local Similarity 100.0%; Pred. No. 8.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 ALLGVLL 18
DB 186 ALLGVLL 193
RESULT 6
ZDHL_MOUSE
ID_ZDHL_MOUSE STANDARD; PRT; 485 AA.
AC Q8RON9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger DHHC domain containing protein 1 (Fragment).
GN ZDHHC1.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Colon;
 RC MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haich F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ussidi T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loughran N.A., Peters G.J., Abramson R.D., Molligh S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gnaratine P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP TISSUE SPECIFICITY.
 RC TISSUE=Pancreas;
 RC MEDLINE=9931009; PubMed=10395086;
 RA Pullina T., Wong P., Gentleman S.;
 RT "The DHHC domain: a new highly conserved cysteine-rich motif.";
 RL Mol. Cell. Biochem. 195:219-226(1999).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- TISSUE SPECIFICITY: Expressed at high levels in fetal lung and
 CC heart. Expressed at lower levels in fetal liver and brain. Also
 CC detected in adult islet cells of pancreas, Leydig cells of testis,
 CC retina and molecular layer of cerebellum.
 CC -1- SIMILARITY: Contains 1 DHHC-type zinc finger.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: BC026570; AAH26570.1; -;
 CC InterPro: IPR001594; Znf.DHHC.
 DR Pfam: PF01529; Zf-DHHC.1.
 DR ProDom: PD003041; Znf.DHHC.1.
 DR PROSITE: PS50216; ZF-DHHC; 1.
 KW Transmembrane; Zinc-finger.
 FT FT NON TER 1 1
 FT TRANSMEM 51 71 POTENTIAL.
 FT TRANSMEM 76 96 POTENTIAL.
 FT TRANSMEM 184 204 POTENTIAL.
 FT TRANSMEM 240 260 POTENTIAL.
 FT ZN FING 132 182 DHHC-TYPE.
 SO SEQUENCE 485 AA; 53107 MW; 866B652C90A49139 CRC64;

Query Match 1.0%; Score 8; DB 1; Length 485;
 Best Local Similarity 100.0%; Pred. No. 8.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ALLGVLL 18
 DB 184 ALLGVLL 191

RESULT 7
 ID USHA_SALPU STANDARD; PRT; 550 AA.
 AC Q9RN37;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DE Protein usha precursor [includes: UDP-sugar hydrolase (EC 3.6.1.45)
 DE (UDP-sugar diphosphatase) (UDP-sugar pyrophosphatase), 5'-nucleotidase
 DE (EC 3.1.3.5) (5'-NT)].
 GN USHA.
 OS Salmonella pullorum.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=605;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21322714; PubMed=11429465;
 RA Innes D., Beacham I.R., Beven C.-A., Douglas M., Laird M.W.,
 RA Joly J.C., Burns D.M.;
 RT "The cryptic usha gene (ushA(c)) in natural isolates of Salmonella
 RT enterica (serotype Typhimurium) has been inactivated by a single
 RT missense mutation.";
 RL Microbiology 147:1867-1896(2001).
 CC -1- FUNCTION: DEGRADATION OF EXTERNAL UDP-GLUCOSE TO URIDINE
 CC MONOPHOSPHATE AND GLUCOSE-1-PHOSPHATE, WHICH CAN THEN BE USED BY
 CC THE CELL (By similarity).
 CC -1- CATALYTIC ACTIVITY: UDP-sugar + H(2)O = UMP + sugar 1-phosphate.
 CC -1- CATALYTIC ACTIVITY: A 5'-ribonucleotide + H(2)O = a ribonucleoside
 CC + phosphate.
 CC -1- COFACTOR: BINDS TWO ZINC IONS (By similarity).
 CC -1- SUBUNIT: Monomer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Periplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE 5'-NUCLEOTIDASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF188727; AAF05581.1; -;
 CC HSP: P07024; IUSH.
 DR InterPro: IPR006179; 5_nucleotidase.
 DR InterPro: IPR006146; 5_NUCLEOTIDASE_2.
 DR InterPro: IPR004843; M-peptidase.
 DR Pfam: PF02872; 5_nucleotidasec.1.
 DR Pfam: PF00149; Metallophos.1.
 DR PRINTS: PR01607; APYRASEFAMILY.
 DR PROSITE: PS00785; 5_NUCLEOTIDASE_1; 1.
 DR PROSITE: PS00786; 5_NUCLEOTIDASE_2; 1.
 KW Hydrolase; Zinc; Periplasmic; Signal; Multifunctional enzyme.
 FT FT CHAIN 1 25
 FT SIGNAL 26 550
 FT ACT SITE 117 117
 FT ACT SITE 120 120
 FT METAL 41 41
 FT METAL 43 43
 FT METAL 84 84
 FT METAL 116 116
 FT METAL 217 217
 FT METAL 252 252
 FT METAL 254 254
 SO SEQUENCE 550 AA; 60604 MW; 655BC78F0D4F62F CRC64;

Query Match 1.0%; Score 8; DB 1; Length 550;
 Best Local Similarity 100.0%; Pred. No. 9.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 675 GGGVLLS 682
 |||||


```

RESULT 10
R7A.DICDI STANDARD; PRT; 78 AA.
AC P14797;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE 40S ribosomal protein S27a.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxId=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89352609; PubMed=2548604;
RA Ohmachi T., Giorda R., Shaw D.R., Ennis H.L.;
RT "Molecular organization of developmentally regulated Dictyostelium
discoidium ubiquitin cDNAs."
RL Biochemistry 28:5226-5231(1989).
CC -1- MISCELLANEOUS: THIS RIBOSOMAL PROTEIN IS SYNTHESIZED AS A
C-TERMINAL EXTENSION PROTEIN (CEP) OF UBIQUITIN.
CC -1- SIMILARITY: BELONGS TO THE S27AE FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch.)
CC -----
DR EMBL, M23750; AAA33264.1; ALT_INIT.
DR DickeyDb; DD01026;
DR InterPro; IPR002906; Ribosomal_S27.
DR Pfam; PF01599; Ribosomal_S27.1.
KW Ribosomal protein; zinc-finger; Metal-binding.
FT DOMAIN 1 24 LYS-RICH (HIGHLY BASIC).
FT ZN_FING 46 69 C4-TYPE.
SQ SEQUENCE 78 AA; 8915 MW; F77B3E09D1B7246F CRC64;

Query Match 0.8%; Score 7; DB 1; Length 78;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 692 KKKKKKT 698
Db 3 KKKKKKT 9

RESULT 11
SLIB.RAT STANDARD; PRT; 104 AA.
AC P09916;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Somatoliberin precursor (Growth hormone-releasing factor) (GRF)
DE (Growth hormone-releasing hormone) (GHRH).
GN GHRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85163768; PubMed=3920534;
RA Mayo K.E., Cerelli G.M., Rosenfeld M.G., Evans R.M.;
RT "Characterization of cDNA and genomic clones encoding the precursor
to rat hypothalamic growth hormone-releasing factor."
RL Nature 314:464-467(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92020929; PubMed=1924334;
RA Gonzalez-Crespo S., Boronat A.;

```

```

RT "Expression of the rat growth hormone-releasing hormone gene in
RT placenta is directed by an alternative promoter."
RT Proc. Natl. Acad. Sci. U.S.A. 88:8749-8753(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Testis;
RX MEDLINE=95203210; PubMed=7895659;
RA Sivasubava C.H., Montes B.S., Rothrock J.K., Peredo M.J.,
RA Pescovitz O.H.;
RT "Presence of a spermatogenic-specific promoter in the rat growth
RT hormone-releasing hormone gene."
RL Endocrinology 136:1502-1508(1995).
RN [4]
RP SEQUENCE OF 31-73.
RC TISSUE=Hypothalamus;
RX MEDLINE=83219259; PubMed=6406907;
RA Spies J., Rivier J., Vale W.;
RT "Characterization of rat hypothalamic growth hormone-releasing
RT factor."
RL Nature 303:532-535(1983).
CC -1- FUNCTION: GRF IS RELEASED BY THE HYPOTHALAMUS AND ACTS ON THE
CC ADENOHYPOPHYSIS TO STIMULATE THE SECRETION OF GROWTH HORMONE.
CC -1- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch.)
CC -----
DR EMBL, X02319; NOT_ANNOTATED_CDS.
DR EMBL, X02335; CAA26194.1;
DR EMBL, X02320; CAA26194.1; JOINED.
DR EMBL, X02321; CAA26194.1; JOINED.
DR EMBL, X02322; CAA26194.1; JOINED.
DR EMBL, M73486; AAA41220.1;
DR EMBL, U01566; AAC52184.1;
DR PIR; A32731; A32731.
DR InterPro; IPR000532; Glucagon.
DR Pfam; PF00123; hormone2; 1.
DR SMART; SM00070; GLUCN.1.
DR PROSITE; PS00260; GLUCAGON.1.
KW Glucagon family; Signal; Hypothalamus.
FT SIGNAL 1 19
FT PEPTIDE 31 73 SOMATOLIBERIN.
SQ SEQUENCE 104 AA; 12266 MW; F9C17485742B2887 CRC64;

Query Match 0.8%; Score 7; DB 1; Length 104;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 387 LTLTSGS 393
Db 10 LTLTSGS 16

RESULT 12
SNPA_STRCS STANDARD; PRT; 132 AA.
AC P56406;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Extracellular small neutral protease (EC 3.4.24.77) (Snapsalysin)
DE (SNCP).
GN SNPA.
OS Streptomyces caespitosus.
OC Bacteria; Actinobacteriae; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxId=53502;
RN [1]

```


RP SEQUENCE.
 RA MEDLINE=96067714; PubMed=7588817;
 RX Harada S., Kinoshita T., Kasai N., Tsunawawa S., Sakiyama F.;
 RT "Complete amino acid sequence of a zinc metalloendoprotease from
 RT Streptomyces caespitosus.";
 RL Eur. J. Biochem. 233:683-686(1995).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
 RX MEDLINE=97244628; PubMed=9089404;
 RA Kurisu G., Kinoshita T., Sugimoto A., Nagara A., Kai Y., Kasai N.,
 RA Harada S.;
 RT "Structure of the zinc endoprotease from Streptomyces caespitosus.";
 RL J. Biochem. 121:304-308(1997).
 CC -1- FUNCTION: SPECIFICALLY HYDROLYZES THE PEPTIDE BOND AT THE IMINO
 CC SIDE OF AROMATIC RESIDUES.
 CC -1- CATALYTIC ACTIVITY: Hydrolyzes proteins with a preference for Tyr
 CC or Phe in the P1' position. Has no action on amino-acid p-
 CC nitroanilides.
 CC -1- COFACTOR: BINDS 1 ZINC ION.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M7.
 DR PIR: S63978; S63978.
 DR PDB: 1KUH; 12-MAR-97.
 DR PDB: 1C7K; 25-APR-01.
 DR MEROPS: M07.001; -;
 DR InterPro: IPR000013; Peptidase_M7.
 DR InterPro: IPR006025; Zn_MTPeptidase.
 DR Pfam: PF02031; Peptidase_M7.1.
 DR PRINTS: PR00787; NEUTRALPTASE.
 DR PRODOM: PD016028; Peptidase_M7.1.
 DR PROSITE: PS00142; ZINC_PROTASE; FALSE_NEG.
 DR Hydroxylase; Metalloprotease; Zinc; 3D-structure.
 FT METAL 83 83 ZINC (CATALYTIC).
 FT ACT SITE 84 84
 FT METAL 87 87 ZINC (CATALYTIC).
 FT METAL 93 93 ZINC (CATALYTIC).
 FT DISULFID 112 112
 FT STRAND 2 10
 FT HELIX 12 14
 FT HELIX 15 26
 FT STRAND 32 38
 FT STRAND 42 47
 FT TURN 50 51
 FT STRAND 54 57
 FT STRAND 63 68
 FT STRAND 69 74
 FT HELIX 77 89
 FT HELIX 90 90
 FT TURN 95 96
 FT TURN 99 100
 FT HELIX 102 104
 FT TURN 105 108
 FT TURN 110 111
 FT HELIX 119 129
 FT TURN 130 132
 SQ SEQUENCE 132 AA; 14376 MW; 7CB988AFC2F0B1E4 CRC64;

Query Match 0.8%; Score 7; DB 1; Length 132;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 521 APPFOQE 527
 DB 11 APPFOQE 17

RESULT 13
 PSBR ARATH STANDARD; PRT; 140 AA.
 AC P27202;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Photosystem II 10 kDa polypeptide, chloroplast precursor.
 GN PSBR OR ART1 OR AT1G79040 OR YUP8H12R.29 OR YUP8H12R.34.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosidia II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=91355945; PubMed=1884004;
 RA Gil-Gomez G., Marrero P.F., Haro D., Ayte J., Hegardt F.G.;
 RT "Characterization of the gene encoding the 10 kDa polypeptide of
 RT photosystem II from Arabidopsis thaliana.";
 RL Plant Mol. Biol. 17:517-522(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Raynal M., Greillet F., Laudie M., Meyer Y., Cooke R., Delzeny M.;
 RL Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Alcafi H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Cressy T.H., Dewar K.,
 RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Maritali A.,
 RA Miltscher J., Miranda M., Nguyen M., Nieman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Ulteback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 408:816-820(2000).
 CC -1- FUNCTION: ASSOCIATED WITH THE OXYGEN-EVOLVING COMPLEX OF
 CC PHOTOSYSTEM II.
 CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane; associated
 CC with the photosystem II complex.
 CC -1- SIMILARITY: BELONGS TO THE PSBR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X55970; CAA39441.1; -;
 DR EMBL: Z17693; CAA79037.1; -;
 DR EMBL: Z17592; CAA79006.1; -;
 DR EMBL: AC002986; AAC17052.1; -;
 DR PIR: S17430; F2MT10.
 DR InterPro: IPR006814; Psbr.
 DR Pfam: PF04725; Psbr.1.
 DR Photosynthesis; Photosystem II; Chloroplast; Transit peptide;
 KW Thylakoid; Membrane.
 FT TRANSIT 1 41 CHLOROPLAST.
 FT CHAIN 42 140 PHOTOSYSTEM II 10 kDa POLYPEPTIDE.
 SQ SEQUENCE 140 AA; 14586 MW; ED41DC7E581637EF CRC64;

Query Match 0.8%; Score 7; DB 1; Length 140;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 566 RGLPSLT 572
 DB 25 RGLPSLT 31

RESULT 14
 PSBR BRACM STANDARD; PRT; 141 AA.

AC P49108;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Photosystem II 10 kDa polypeptide, chloroplast precursor.
 GN PSBR.

OS Brassica campestris (Field mustard).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC euroside II; Brassicales; Brassicaceae; Brassica.
 OC NCBI_TaxId=3711;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Pekinensis;
 RX MEDLINE=95334521; PubMed=7610192;
 RA Kim H.U., Yun C.H., Park B.S., Ryu J.C., Chung T.Y.;
 RT "Nucleotide sequence of a cDNA clone encoding the complete precursor
 for the '10-kilodalton' polypeptide of photosystem II from Chinese
 cabbage.";
 RL Plant Physiol. 108:867-867(1995).
 CC -1- FUNCTION: ASSOCIATED WITH THE OXYGEN-EVOLVING COMPLEX OF
 PHOTOSYSTEM II.
 CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane; associated
 CC with the photosystem II complex.
 CC -1- SIMILARITY: BELONGS TO THE PSBR FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: L31936; AAA74957.1; -
 DR InterPro: IPR006814; Psbr.
 DR Pfam: PF04725; Psbr; 1.
 DR Photosynthesis; Photosystem II; Chloroplast; Transit peptide;
 KW Thylakoid; Membrane.
 FT TRANSIT 1 42 CHLOROPLAST (BY SIMILARITY).
 FT CHAIN 43 141 PHOTOSYSTEM II 10 KDA POLYPEPTIDE.
 SQ SEQUENCE 141 AA; 14649 MW; AA85D2561ECFEFA CRC64;

Query Match 0.8%; Score 7; DB 1; Length 141;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 566 RGLPSLT 572
 DB 25 RGLPSLT 31

RESULT 15
 MOAC METUA STANDARD; PRT; 152 AA.

AC Q58535;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable molybdenum cofactor biosynthesis protein C.
 GN MOAC OR MJ1135.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococcales; Methanocaldococcus.

OK NCBI_TaxId=2190;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8668087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Claydon R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
 RA Uiterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Upton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Frazer C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii.";
 RL Science 273:1058-1073(1996).
 CC -1- FUNCTION: Together with moaA, is involved in the conversion of a
 CC guanosine derivative (GXP) into molybdopterin precursor Z (By
 CC similarity).
 CC -1- PATHWAY: Molybdenum cofactor biosynthesis; first step.
 CC -1- SIMILARITY: Belongs to the moac family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: U67556; AAB99137.1; -
 DR PIR: F64441; F64441.
 DR HSSP: P30747; 1EKR.
 DR TIGR: MJ1135; -
 DR HAMAP: MF_01224; -; 1.
 DR InterPro: IPR002820; Moac.
 DR Pfam: PF01967; Moac; 1.
 DR TIGRFAMs: TIGR00581; moac; 1.
 KW Molybdenum cofactor biosynthesis; Complete proteome.
 FT ACT SITE 123 POTENTIAL.
 SQ SEQUENCE 152 AA; 17043 MW; D91ACGED1D2852AF CRC64;

Query Match 0.8%; Score 7; DB 1; Length 152;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GVSIALL 13
 DB 113 GVSIALL 119

Search completed: February 20, 2004, 19:47:10
 Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: February 20, 2004, 18:39:32 ; Search time 119 Seconds
(without alignments)
1812.874 Million cell updates/sec

Title: US-09-899-569a-4

Perfect score: 836
Sequence: 1 MAGLNCGVSIALLGVLLGA.....SSKDTDPLNTQEPMPAP 836

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 830525 seags, 258052604 residues

Word size: 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mmc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	836	100.0	836	4	Q96GQ7
2	725	86.7	836	4	Q9H5V8
3	382	45.7	649	4	Q9H8C2
4	341	40.8	343	4	Q8WU91
5	291	34.8	392	4	Q9H676
6	39	4.7	320	11	Q8K246
7	39	4.7	470	11	Q8BZD9
8	27	3.2	384	11	Q921W9
9	9	1.1	108	5	Q96238
10	9	1.1	815	5	Q23860
11	9	1.1	1245	5	Q812R3
12	9	1.1	1629	5	Q9U0X9
13	9	1.1	2747	5	Q81C5
14	8	1.0	108	11	Q8BR25
15	8	1.0	127	11	Q8C3C4
16	8	1.0	144	10	Q9LEZ1

17	8	1.0	173	8	Q8HK3	Q8HK3 arcs sp. k
18	8	1.0	179	4	Q9UHQ2	Q9UHQ2 homo sapien
19	8	1.0	187	16	Q8X6W2	Q8X6W2 escherichia
20	8	1.0	190	4	Q96S03	Q96S03 homo sapien
21	8	1.0	195	16	Q8CVR9	Q8CVR9 escherichia
22	8	1.0	203	5	Q17789	Q17789 caenorhabdi
23	8	1.0	205	17	Q8TVN0	Q8TVN0 methanopyru
24	8	1.0	206	16	Q8DM11	Q8DM11 synechococc
25	8	1.0	221	4	Q9BRV3	Q9BRV3 homo sapien
26	8	1.0	221	4	Q9UHQ3	Q9UHQ3 homo sapien
27	8	1.0	221	6	Q95XW8	Q95XW8 papio anubi
28	8	1.0	232	16	Q8ZS89	Q8ZS89 anabena sp
29	8	1.0	262	16	Q9KD01	Q9KD01 bacillus ha
30	8	1.0	285	5	Q9U632	Q9U632 trichomegal
31	8	1.0	349	5	Q9BIJ3	Q9BIJ3 sarcocystis
32	8	1.0	380	10	Q82468	Q82468 mesembryant
33	8	1.0	390	3	Q9USL5	Q9USL5 schizosach
34	8	1.0	443	2	Q8RTV3	Q8RTV3 uncultured
35	8	1.0	468	3	Q9P3W6	Q9P3W6 schizosach
36	8	1.0	484	5	Q9VAF4	Q9VAF4 drosophila
37	8	1.0	484	11	Q8BJ24	Q8BJ24 mus musculu
38	8	1.0	509	5	Q9BID1	Q9BID1 anophles g
39	8	1.0	510	10	Q8W2M4	Q8W2M4 nicotiana t
40	8	1.0	550	2	Q9RN37	Q9RN37 salmoneila
41	8	1.0	565	5	Q81151	Q81151 plasmodium
42	8	1.0	597	11	Q9JKC1	Q9JKC1 ratius norv
43	8	1.0	682	16	Q9HYO5	Q9HYO5 pseudomonas
44	8	1.0	899	5	Q97296	Q97296 plasmodium
45	8	1.0	1137	11	Q8K561	Q8K561 mus musculu

ALIGNMENTS

RESULT 1

ID	Q96GQ7	PRELIMINARY;	PRT;	836 AA.
AC	Q96GQ7;			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	CUB domain containing protein 1.			
GN	CDCP1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21359860; PubMed=11466621;			
RA	Scherl-Mostaguer M., Sommergruber W., Abseher R., Hauptmann R.,			
RA	Ambros P., Schweifer N.;			
RT	"Identification of a novel gene, CDCP1, overexpressed in human			
RT	colorectal cancer."			
RL	Oncogene 20:4402-4408(2001).			
DR	EMBL: AY026461; AAK02058.1; -			
SQ	SEQUENCE 836 AA; 92873 MW; FB4D2DBDD35C519 CRC64;			

Query Match	100.0%; Score 836; DB 4; Length 836;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 836; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MAGLNCGVSIALLGVLLGAARLPFGAARFIALPRESNITVLYKGPPTLLAKCYVI 60
DB	1 MAGLNCGVSIALLGVLLGAARLPFGAARFIALPRESNITVLYKGPPTLLAKCYVI 60
QY	61 SKRHITMISIGERIVTFSCQSPENHFVIEIQNIDCMGSPCFGVOLOPSTSLPT 120
DB	61 SKRHITMISIGERIVTFSCQSPENHFVIEIQNIDCMGSPCFGVOLOPSTSLPT 120
QY	121 LNRFTIWDVKAHKSIGLELOPSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB	121 LNRFTIWDVKAHKSIGLELOPSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180

```
QY 181 GTVSRIKMOEGYKMAHLPMFHPNRVSGFSIANRSSIKRLCTIESEVEGEGSATLMSANY 240
DB 181 GTVSRIKMOEGYKMAHLPMFHPNRVSGFSIANRSSIKRLCTIESEVEGEGSATLMSANY 240
QY 241 PGCFPEDELMTMOQFVVPAPHLRASVSFLNPNLSNCRKEERVEYYIIGSTTNEVFKLEBK 300
DB 241 PGCFPEDELMTMOQFVVPAPHLRASVSFLNPNLSNCRKEERVEYYIIGSTTNEVFKLEBK 300
QY 301 QPGNMGNFNLSLQGCDDAOSPGILRLQFVLYQHPQNESNKIYVVDLSNERAMSLTIE 360
DB 301 QPGNMGNFNLSLQGCDDAOSPGILRLQFVLYQHPQNESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVQSRKFPVPGCFVCLSESRCTSSNLTLTSGSKHISFLCDLTRLMMNVEKTIISCTDH 420
DB 361 PRPVQSRKFPVPGCFVCLSESRCTSSNLTLTSGSKHISFLCDLTRLMMNVEKTIISCTDH 420
QY 421 RYCQRKSYSLQVPSDILHLPVELHDFSMKLVKPKRLSLVLPKQLQOHTHEKPCNTSF 480
DB 421 RYCQRKSYSLQVPSDILHLPVELHDFSMKLVKPKRLSLVLPKQLQOHTHEKPCNTSF 480
QY 481 SYLVASAIPSODLYGSGFCPGGSIKOIYQKONI SVTLRTFAPSFOQEASRQGLTYSFIY 540
DB 481 SYLVASAIPSODLYGSGFCPGGSIKOIYQKONI SVTLRTFAPSFOQEASRQGLTYSFIY 540
QY 541 FKEBGVFTVPTDKSKVYLRTPNMDRGLPSLTSVSNISVPRDQVACLTFEKRSGVVCQ 600
DB 541 FKEBGVFTVPTDKSKVYLRTPNMDRGLPSLTSVSNISVPRDQVACLTFEKRSGVVCQ 600
QY 601 TGRAPMIIOEORTRAEISLDEVDLPKPSFHHHSFWNINSCSPISGQDLDFSVLIT 660
DB 601 TGRAPMIIOEORTRAEISLDEVDLPKPSFHHHSFWNINSCSPISGQDLDFSVLIT 660
QY 661 PRVDLTVILIAAVGGVLLSALGLIICVKKKKKTKNGPAVGYNGNINTEMPROK 720
DB 661 PRVDLTVILIAAVGGVLLSALGLIICVKKKKKTKNGPAVGYNGNINTEMPROK 720
QY 721 KFQGRKNDSHVAVIEDTMVYGHLLQDSSGSLQPEVDIYRPPQGTMGVCPSPPTIC 780
DB 721 KFQGRKNDSHVAVIEDTMVYGHLLQDSSGSLQPEVDIYRPPQGTMGVCPSPPTIC 780
QY 781 SRAPFAKLATEEPPRSPSESESEPTFSHPNNGVSSKOTDIPLINTOEPMEPAE 836
DB 781 SRAPFAKLATEEPPRSPSESESEPTFSHPNNGVSSKOTDIPLINTOEPMEPAE 836

RESULT 2
Q9H5V8 PRELIMINARY; PRT; 836 AA.
ID Q9H5V8 AC Q9H5V8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ22969.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OC NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Ohtsuka M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Iwagaki T., Sugano S.,
RT "NBO human cDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK026622; BAB1511.1; -.
KW Hypothetical protein.
SQ SEQUENCE 836 AA; 92874 MW; 9B980475C3E5C4C8 CRC64;
```

```
Query Match 86.7%; Score 725; DB 4; Length 836;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 825; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 MAGLNCGVSIALLGVLLIGAAARLPRGAFAFIALPRESENTIVYLKGTPTLLAKCYIYI 60
DB 1 MAGLNCGVSIALLGVLLIGAAARLPRGAFAFIALPRESENTIVYLKGTPTLLAKCYIYI 60
QY 61 SKRHITMLSIKSGEIVTFPGSQSPENHFVIEIQKIDCMGSPCFGEVQLOPSTSLPT 120
DB 61 SKRHITMLSIKSGEIVTFPGSQSPENHFVIEIQKIDCMGSPCFGEVQLOPSTSLPT 120
QY 121 LNRTPIWDVKAHKSIGLELQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCGN 180
DB 121 LNRTPIWDVKAHKSIGLELQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCGN 180
QY 181 GTVSRIKMOEGYKMAHLPMFHPNRVSGFSIANRSSIKRLCTIESEVEGEGSATLMSANY 240
DB 181 GTVSRIKMOEGYKMAHLPMFHPNRVSGFSIANRSSIKRLCTIESEVEGEGSATLMSANY 240
QY 241 PGCFPEDELMTMOQFVVPAPHLRASVSFLNPNLSNCRKEERVEYYIIGSTTNEVFKLEBK 300
DB 241 PGCFPEDELMTMOQFVVPAPHLRASVSFLNPNLSNCRKEERVEYYIIGSTTNEVFKLEBK 300
QY 301 QPGNMGNFNLSLQGCDDAOSPGILRLQFVLYQHPQNESNKIYVVDLSNERAMSLTIE 360
DB 301 QPGNMGNFNLSLQGCDDAOSPGILRLQFVLYQHPQNESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVQSRKFPVPGCFVCLSESRCTSSNLTLTSGSKHISFLCDLTRLMMNVEKTIISCTDH 420
DB 361 PRPVQSRKFPVPGCFVCLSESRCTSSNLTLTSGSKHISFLCDLTRLMMNVEKTIISCTDH 420
QY 421 RYCQRKSYSLQVPSDILHLPVELHDFSMKLVKPKRLSLVLPKQLQOHTHEKPCNTSF 480
DB 421 RYCQRKSYSLQVPSDILHLPVELHDFSMKLVKPKRLSLVLPKQLQOHTHEKPCNTSF 480
QY 481 SYLVASAIPSODLYGSGFCPGGSIKOIYQKONI SVTLRTFAPSFOQEASRQGLTYSFIY 540
DB 481 SYLVASAIPSODLYGSGFCPGGSIKOIYQKONI SVTLRTFAPSFOQEASRQGLTYSFIY 540
QY 541 FKEBGVFTVPTDKSKVYLRTPNMDRGLPSLTSVSNISVPRDQVACLTFEKRSGVVCQ 600
DB 541 FKEBGVFTVPTDKSKVYLRTPNMDRGLPSLTSVSNISVPRDQVACLTFEKRSGVVCQ 600
QY 601 TGRAPMIIOEORTRAEISLDEVDLPKPSFHHHSFWNINSCSPISGQDLDFSVLIT 660
DB 601 TGRAPMIIOEORTRAEISLDEVDLPKPSFHHHSFWNINSCSPISGQDLDFSVLIT 660
QY 661 PRVDLTVILIAAVGGVLLSALGLIICVKKKKKTKNGPAVGYNGNINTEMPROK 720
DB 661 PRVDLTVILIAAVGGVLLSALGLIICVKKKKKTKNGPAVGYNGNINTEMPROK 720
QY 721 KFQGRKNDSHVAVIEDTMVYGHLLQDSSGSLQPEVDIYRPPQGTMGVCPSPPTIC 780
DB 721 KFQGRKNDSHVAVIEDTMVYGHLLQDSSGSLQPEVDIYRPPQGTMGVCPSPPTIC 780
QY 781 SRAPFAKLATEEPPRSPSESESEPTFSHPNNGVSSKOTDIPLINTOEPMEPAE 836
DB 781 SRAPFAKLATEEPPRSPSESESEPTFSHPNNGVSSKOTDIPLINTOEPMEPAE 836

RESULT 3
Q9H8C2 PRELIMINARY; PRT; 649 AA.
ID Q9H8C2 AC Q9H8C2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ13772.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OC NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Iwagaki T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
```

RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA Tanase T., Nomura Y., Togiyasu S., Komai F., Hara R., Takeuchi K.,
RA Arima M., Nabebara T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA Makamatsu A., Nakamura Y., Nagahara K., Masuo Y., Oshima A.,
RT "NEDO human cDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AK023834; BAB14695.1; -
KW Hypothetical protein.
SQ SEQUENCE 649 AA; 72579 MW; 0F4404AD31D60E9A CRC64;

Query Match 45.7%; Score 382; DB 4; Length 649;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 582; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 253 QFVPAHLRASVSLFNLNSCERKEEVEYIIPGTTNPVEFKEDKQPMAGNPLS 312
DB 66 QFVPAHLRASVSLFNLNSCERKEEVEYIIPGTTNPVEFKEDKQPMAGNPLS 125
QY 313 LOGCDQDASPGILRLQFQVLYVHPQNESNKIYVVDLSNERAMSITTEPRPVKSRKEVP 372
DB 126 LOGCDQDASPGILRLQFQVLYVHPQNESNKIYVVDLSNERAMSITTEPRPVKSRKEVP 185
QY 373 GCFPLCSRTSSNLTLSGSKHISFLCDLTLRMNVEKTSCTDHRVYCKRSYSLOV 432
DB 186 GCFPLCSRTSSNLTLSGSKHISFLCDLTLRMNVEKTSCTDHRVYCKRSYSLOV 245
QY 433 PSDILHLPEVLHDSWKLVPKDRSLVLPAPKQLQHTHEKPCNTSPSYLVASIPSD 492
DB 246 PGDILHLPEVLHDSWKLVPKDRSLVLPAPKQLQHTHEKPCNTSPSYLVASIPSD 305.
QY 493 LYFGSFCGSGIKQIQVKNISVTLRTAPSPFOEASRQGLTVSEIPIYKEGVEFTVPD 552
DB 306 LYFGSFCGSGIKQIQVKNISVTLRTAPSPFOEASRQGLTVSEIPIYKEGVEFTVPD 365
QY 553 TKSRYLRTPMWDGSLPGLSVSNISVPRDQVACLTFKRSQVYCTGAPMIOQR 612
DB 366 TKSRYLRTPMWDGSLPGLSVSNISVPRDQVACLTFKRSQVYCTGAPMIOQR 425
QY 613 TRAEIISLDEVDLPKPSFHHSFVWVNSCSPSGKQDLIFSUTLTPRTVDLTVLIA 672
DB 426 TRAEIISLDEVDLPKPSFHHSFVWVNSCSPSGKQDLIFSUTLTPRTVDLTVLIA 485
QY 673 AVGGGVLISALGLIICVKKKKKTKNGPAVINGNINTEMPROPKPKQGRKNDSH 732
DB 486 AVGGGVLISALGLIICVKKKKKTKNGPAVINGNINTEMPROPKPKQGRKNDSH 545
QY 733 VYAVIEDTMVYGHLLQDSSGSLPQEVDTYRPPQGTWGVCPSPPTICSRAPLATEE 792
DB 546 VYAVIEDTMVYGHLLQDSSGSLPQEVDTYRPPQGTWGVCPSPPTICSRAPLATEE 605
QY 793 PPRSPSESEPTESHNNNDVSSKQTDIPLNTQEPMEPAE 836
DB 606 PPRSPSESEPTESHNNNDVSSKQTDIPLNTQEPMEPAE 649

RESULT 4

Q8WU91 PRELIMINARY; PRT: 343 AA.
AC Q8WU91;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Similar to hypothetical protein FLJ22969.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strauberg R.;
RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC021099; AAH21099.1; -

KW Hypothetical protein.
SQ SEQUENCE 343 AA; 37817 MW; 3E4E13379D9D94D1B CRC64;

Query Match 40.8%; Score 341; DB 4; Length 343;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSLILGLVLLGAARLPRGAFAFIALPRESNITVLKIGPTLLAKPCYVI 60
DB 1 MAGNCGVSLILGLVLLGAARLPRGAFAFIALPRESNITVLKIGPTLLAKPCYVI 60
QY 61 SKRIITMLISGRIYFTFSCQSPENHVEIIOKNIDCMGCPGPGVQLOPSTSLPT 120
DB 61 SKRIITMLISGRIYFTFSCQSPENHVEIIOKNIDCMGCPGPGVQLOPSTSLPT 120
QY 121 LNRFTFMDVKAHKSIGLEQFSLRLQIGBESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTFMDVKAHKSIGLEQFSLRLQIGBESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIMQEGVKAALHLPFHPNVSQFSIANSISIKLCTISVFEGBSATLMSANY 240
DB 181 GTVSRIMQEGVKAALHLPFHPNVSQFSIANSISIKLCTISVFEGBSATLMSANY 240
QY 241 PEGFPEDELMTWQFVPAHLRASVSLFNLNSCERKEEVEYIIPGTTNPVEFKEDK 300
DB 241 PEGFPEDELMTWQFVPAHLRASVSLFNLNSCERKEEVEYIIPGTTNPVEFKEDK 300
QY 301 QPGMAGNPLSLQCDQDASPGILRLQFQVLYVHPQNES 341
DB 301 QPGMAGNPLSLQCDQDASPGILRLQFQVLYVHPQNES 341

RESULT 5

Q9H676 PRELIMINARY; PRT: 392 AA.
AC Q9H676;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)
DE Hypothetical protein FLJ22534 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Oktani R., Ota T., Suzuki Y., Odayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isegai T., Sugano S.;
RT "NEDO human cDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AK026187; BAB15388.1; -
KW Hypothetical protein.
FT NON TER
SQ SEQUENCE 392 AA; 44031 MW; 3E21C8B2703E52F3 CRC64;

Query Match 34.8%; Score 291; DB 4; Length 392;
Best Local Similarity 99.7%; Pred. No. 6.4e-290;
Matches 391; Conservative 1; Indels 0; Gaps 0;

QY 305 MAGNPLSLQCDQDASPGILRLQFQVLYVHPQNESNKIYVVDLSNERAMSITTEPRPV 364
DB 1 MAGNPLSLQCDQDASPGILRLQFQVLYVHPQNESNKIYVVDLSNERAMSITTEPRPV 60
QY 365 KQSRKFPVPGCFVCLSRSSNLTLSGSKHISFLCDLTLRMNVEKTSCTDHRVYCO 424
DB 61 KQSRKFPVPGCFVCLSRSSNLTLSGSKHISFLCDLTLRMNVEKTSCTDHRVYCO 120
QY 425 RKYSYLOVPSDILHLPEVLHDSWKLVPKDRSLVLPAPKQLQHTHEKPCNTSPSYLV 484
DB 121 RKYSYLOVPSDILHLPEVLHDSWKLVPKDRSLVLPAPKQLQHTHEKPCNTSPSYLV 180
QY 485 ASAIPQDLVFGSFCGSGIKQIQVKNISVTLRTAPSPFOEASRQGLTVSEIPIYKEE 544

```

DB 181 AALPSQDLYFSGFCGGSIKQIQVKNQNTSVTLRTFAPSFQOASRQGLTVSFIPFKKE 240
OY 545 GVFTVTPDTRSKYVLTPTMWRGRLPSLTSVSNVISVPRQVACTFFPKERSGVVCOTGRA 604
DB 241 GVFTVTPDTRSKYVLTPTMWRGRLPSLTSVSNVISVPRQVACTFFPKERSGVVCOTGRA 300
OY 605 FMIOEORTABEISLSDVDLPKPSFHHHSFWNVINSNCSPTSGQDLDFSVTLTPRTV 664
DB 301 FMIOEORTABEISLSDVDLPKPSFHHHSFWNVINSNCSPTSGQDLDFSVTLTPRTV 360
OY 665 DLTVLIAVGGVLLSLGLICVCKKKK 696
DB 361 DLTVLIAVGGVLLSLGLICVCKKKK 392

```

```

RESULT 6
O8K246 PRELIMINARY; PRT; 320 AA.
AC O8K246;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DE 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC034137; AAH34137.1; -
KW Hypothetical protein.
SQ SEQUENCE 320 AA; 35962 MW; 050906547833BC36 CRC64;

```

```

Query Match 4.7%; Score 39; DB 11; Length 320;
Best Local Similarity 100.0%; Pred. No. 5.8e-31;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 297 LEDKOPGNMAGNPNLSLOGCDODASPGILRLQFOVLVQ 335
DB 110 LEDKOPGNMAGNPNLSLOGCDODASPGILRLQFOVLVQ 148

```

```

RESULT 7
O8BZD9 PRELIMINARY; PRT; 470 AA.
AC O8BZD9;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Similar to CUB domain containing protein 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Urinary bladder;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analyses of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AC035674; BAC29147.1; -
SQ SEQUENCE 470 AA; 52277 MW; A36CD2B547348383 CRC64;

```

```

Query Match 4.7%; Score 39; DB 11; Length 470;
Best Local Similarity 100.0%; Pred. No. 8.1e-31;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 297 LEDKOPGNMAGNPNLSLOGCDODASPGILRLQFOVLVQ 335
DB 297 LEDKOPGNMAGNPNLSLOGCDODASPGILRLQFOVLVQ 335

```

```

RESULT 8
O921M9 PRELIMINARY; PRT; 384 AA.
AC O921M9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Unknown (Protein for IMAGE:3709937) (Fragment).
GN 9030022E12R1K OR A4409659.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011340; AAH11340.1; -
DR MGD; MGI:1925928; 9030022E12R1K.
FT NON_TER 1
SQ SEQUENCE 384 AA; 42406 MW; 4B5F17D30FE16080 CRC64;

```

```

Query Match 3.2%; Score 27; DB 11; Length 384;
Best Local Similarity 100.0%; Pred. No. 1.5e-18;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 724 KGRKNDSHVYAVIEDTWYGHLLQDS 750
DB 272 KGRKNDSHVYAVIEDTWYGHLLQDS 298

```

```

RESULT 9
O96238 PRELIMINARY; PRT; 108 AA.
AC O96238;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Plasmodium falciparum (Isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxId=36329;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=3D7;
RX MEDLINE=99021743; PubMed=9804551;
RA Gardner M.J., Tetteh H., Carucci D.J., Cummings L.M., Aravind L.,
RA Koonin E.V., Shallow S., Mason T., Yu K., Fujii C., Pedersen J.,
RA Shen K., Jing J., Ascon C., Lai Z., Schwartz D.C., Petrea M.,
RA Salberg S., Zhou L., Sutton G.G., Clayton R., White H.O.,
RA Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium
RL falciparum.";
RL Science 282:1126-1132(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Nene V., Shallow S.J., Craig A., Kyse S.,
RA Chan M.-S., Nene V., Shallow S.J., Sub B., Peterson J., Angiolini S.,
RA Persea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrett B.;

```

RT "Genome sequence of the human malaria parasite Plasmodium
 RT falciparum.";
 RL Nature 419:498-511(2002).
 DR EMBL; AB001414; AAC71934.2; -
 KW Hypothetical protein.
 SQ SEQUENCE 108 AA; 13675 MW; 11DC06350C0362B5 CRC64;
 Query Match 1.1%; Score 9; DB 5; Length 108;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 692 KKKKKKTKN 700
 DB 100 KKKKKKTKN 108
 RESULT 10
 ID Q23860 PRELIMINARY; PRT; 815 AA.
 AC Q23860;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 20, Last annotation update)
 DE Vacuolar proton ATPase 100-kDa subunit.
 GN VATM.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96338613; PubMed=8743951;
 RA Liu T., Clarke M.;
 RT "The vacuolar proton pump of Dictyostelium discoideum: molecular
 cloning and analysis of the 100 kDa subunit.";
 RL J. Cell Sci. 109:1041-1051(1996).
 DR EMBL; U38803; AAB49621.1; -
 DR InterPro; IPR002490; V_ATPase_sub116.
 DR Pfam; PF01496; V_ATPase_sub_a; 1.
 SQ SEQUENCE 815 AA; 93272 MW; DF0858C79F9F75D2 CRC64;
 Query Match 1.1%; Score 9; DB 5; Length 815;
 Best Local Similarity 100.0%; Pred. No. 9.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 679 LLLSALGLI 687
 DB 430 LLLSALGLI 438
 RESULT 11
 ID O812R3 PRELIMINARY; PRT; 1245 AA.
 AC O812R3;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN PF11200W.
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22255708; PubMed=13368867;
 RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
 RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brookes K.,
 RA Buceae C.O., Burrows C., Cherevach I., Chillingworth C.,
 RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corson C.,
 RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
 RA Feltwell T., Goble A., Goodhead I., Gilliam R., Hamlin N., Hance Z.,
 RA Harper D., Hauser H., Hornsby T., Holtroyd S., Horrocks P.,
 RA Humphrey S., Jagsels K., James K.D., Johnson D., Kerhornou A.,
 RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,

RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
 RA Oliver K., Ormond D., Price C., Quail M.A., Rabinovitch E.,
 RA Rajendram M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
 RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
 RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
 RA Sultson J.E., Craig A., Newbold C., Barrett B.G.;
 RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
 RL Nature 419:527-531(2002).
 DR EMBL; AL929357; CAD51926.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 1245 AA; 144680 MW; 12904C239AED17F1 CRC64;
 Query Match 1.1%; Score 9; DB 5; Length 1245;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 692 KKKKKKTKN 700
 DB 690 KKKKKKTKN 698
 RESULT 12
 ID Q9U0K9 PRELIMINARY; PRT; 1629 AA.
 AC Q9U0K9;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE Hypothetical 197.4 kDa protein.
 GN MAL4P2.13.
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=3D7;
 RA Devlin K., Bowman S., Churcher C., Harris B., Harris D., Lawson D.,
 RA Quail M., Barrett B.;
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL035475; CAB62854.2; -
 KW Hypothetical protein.
 SQ SEQUENCE 1629 AA; 197412 MW; C1233E537648C9D CRC64;
 Query Match 1.1%; Score 9; DB 5; Length 1629;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 692 KKKKKKTKN 700
 DB 498 KKKKKKTKN 506
 RESULT 13
 ID O81CL5 PRELIMINARY; PRT; 2747 AA.
 AC O81CL5;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Ribonuclease, putative.
 GN MAL6P1.272.
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cherevach I., Davis P., Goodhead I., Stevens K., Mungall K.,
 RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,
 RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrett B.,
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL844505; CAD50426.1; -
 SQ SEQUENCE 2747 AA; 325595 MW; F01B08D441D0D3E5 CRC64;

Query Match 1.1%; Score 9; DB 5; Length 2747;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 692 KKKKKTNK 700
Db 934 KKKKKTNK 942

Db 29 KKKKKTNK 36
Search completed: February 20, 2004, 19:49:22
Job time : 123 secs

RESULT 14

O8BR25 PRELIMINARY; PRT; 108 AA.
AC O8BR25;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK045823; BAC32505.1; -.
KW Hypothetical protein.
SQ SEQUENCE 108 AA; 12201 MW; A9F6B7CC972EA55B CRC64;

Query Match 1.0%; Score 8; DB 11; Length 108;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 693 KKKKKTNK 700
Db 10 KKKKKTNK 17

RESULT 15

O8C3C4 PRELIMINARY; PRT; 127 AA.
ID O8C3C4
AC O8C3C4;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK086308; BAC39646.1; -.
KW Hypothetical protein.
SQ SEQUENCE 127 AA; 14090 MW; 24E2FC1C01F1EA2F CRC64;

Query Match 1.0%; Score 8; DB 11; Length 127;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 693 KKKKKTNK 700

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2004, 19:45:04 ; Search time 32 Seconds
(without alignments)
1105.371 Million cell updates/sec

Title: US-09-899-569A-4
Perfect score: 836
Sequence: 1 MAGLNCGVSIALLGVLLGA.....SSKTDIPLNTQEPMEPAE 836

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents_AA:
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	413	49.4	414	US-09-489-847-273	Sequence 273, App
2	413	49.4	443	US-09-489-847-271	Sequence 271, App
3	211	25.2	709	US-09-489-847-132	Sequence 132, App
4	159	19.0	159	US-09-489-847-272	Sequence 272, App
5	10	1.2	206	US-09-252-991A-18034	Sequence 18034, A
6	9	1.1	19	US-09-333-599-3	Sequence 3, Appl1
7	9	1.1	19	US-09-499-781-3	Sequence 3, Appl1
8	8	1.0	71	US-09-369-247-79	Sequence 79, Appl1
9	8	1.0	156	US-09-252-991A-32450	Sequence 32450, A
10	8	1.0	221	US-09-599-360B-104	Sequence 104, App
11	8	1.0	703	US-09-252-991A-17865	Sequence 17865, A
12	7	0.8	57	US-09-082-358B-43	Sequence 43, Appl1
13	7	0.8	66	US-09-205-258-381	Sequence 381, App
14	7	0.8	98	US-09-732-210-341	Sequence 341, App
15	7	0.8	77	US-08-946-329A-94	Sequence 94, Appl1
16	7	0.8	104	US-09-252-991A-29655	Sequence 29655, A
17	7	0.8	104	US-09-252-991A-18492	Sequence 18492, A
18	7	0.8	105	US-09-328-352-4768	Sequence 4768, App
19	7	0.8	122	US-09-107-532A-6744	Sequence 6744, App
20	7	0.8	136	US-09-328-352-6544	Sequence 6544, App
21	7	0.8	147	US-08-986-837-2	Sequence 2, Appl1
22	7	0.8	147	US-09-809-593-2	Sequence 2, Appl1
23	7	0.8	149	US-09-252-991A-18492	Sequence 18492, A
24	7	0.8	155	US-09-252-991A-32399	Sequence 32399, A
25	7	0.8	187	US-09-252-991A-18109	Sequence 18109, A
26	7	0.8	222	US-09-252-991A-27763	Sequence 27763, A
27	7	0.8	273	US-09-252-991A-20506	Sequence 20506, A

28	7	0.8	278	US-09-252-991A-20931	Sequence 20931, A
29	7	0.8	352	US-09-107-532A-6703	Sequence 6703, App
30	7	0.8	355	US-09-134-001C-5391	Sequence 5391, App
31	7	0.8	359	US-09-252-991A-26124	Sequence 26124, A
32	7	0.8	377	US-09-161-994A-12	Sequence 12, Appl1
33	7	0.8	394	US-09-390-721-5	Sequence 5, Appl1
34	7	0.8	434	US-09-252-991A-23489	Sequence 23489, A
35	7	0.8	436	US-09-252-991A-23267	Sequence 23267, A
36	7	0.8	437	US-09-252-991A-24572	Sequence 24572, A
37	7	0.8	452	US-09-328-352-7162	Sequence 7162, App
38	7	0.8	473	US-09-252-991A-28845	Sequence 28845, A
39	7	0.8	493	US-09-117-250-3	Sequence 3, Appl1
40	7	0.8	499	US-09-252-991A-21491	Sequence 21491, A
41	7	0.8	506	US-09-117-250-1	Sequence 1, Appl1
42	7	0.8	531	US-08-933-750C-9	Sequence 9, Appl1
43	7	0.8	531	US-09-234-613-9	Sequence 9, Appl1
44	7	0.8	534	US-09-252-991A-17757	Sequence 17757, A
45	7	0.8	539	US-09-252-991A-29555	Sequence 29555, A

ALIGNMENTS

```
RESULT 1
US-09-489-847-273
; Sequence 273, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031p1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 273
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-489-847-273
```

Query Match 49.4%; Score 413; DB 4; Length 414;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAGLNCGVSIALLGVLLGAARLPFGAEFAIALPRESNTIVYLKGPPTLAKCYVI	60
DB	1	MAGLNCGVSIALLGVLLGAARLPFGAEFAIALPRESNTIVYLKGPPTLAKCYVI	60
QY	61	SKRHITMISGERTVTFSCQSPENHFVLEIQNDICMSGPCFGEVQLOPSTSLPT	120
DB	61	SKRHITMISGERTVTFSCQSPENHFVLEIQNDICMSGPCFGEVQLOPSTSLPT	120
QY	121	LNRTFTWVKAHKSIGLELOFSIPRLROIGPESCPDVTSHISGRIDATVTRIGTFCSN	180
DB	121	LNRTFTWVKAHKSIGLELOFSIPRLROIGPESCPDVTSHISGRIDATVTRIGTFCSN	180
QY	181	GTVSRIKQEGVKALHPFHPRVNVSFGSTANRSSIRKLCIESVFGEBSATLMSANY	240
DB	181	GTVSRIKQEGVKALHPFHPRVNVSFGSTANRSSIRKLCIESVFGEBSATLMSANY	240

QY 241 PEGFPEDELMTQFVPAHLRASVSFLNRLSNCEKREVEYIIGSTTNPVEFKLEDK 300
DB 241 PEGFPEDELMTQFVPAHLRASVSFLNRLSNCEKREVEYIIGSTTNPVEFKLEDK 300
QY 301 QPNNMAGNFNLSLQGGDDDAQSPGILRLQFOVLVQHPONESNKIYVVDLSNERAMSLTTE 360
DB 301 QPNNMAGNFNLSLQGGDDDAQSPGILRLQFOVLVQHPONESNKIYVVDLSNERAMSLTTE 360
QY 361 PRPVKQSRKRVPCGFCVLESRTCSNLTLTSGSKHKISFLCDDLTRLMNVK 413
DB 361 PRPVKQSRKRVPCGFCVLESRTCSNLTLTSGSKHKISFLCDDLTRLMNVK 413

RESULT 2

US-09-847-271
; Sequence 271, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 271
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-489-847-271

Query Match 49.4%; Score 413; DB 4; Length 443;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCVSIALLGVLLGAAALPRGAFAFEIALPRESNITVLIKGTPTLLAKPCYIYI 60
DB 30 MAGLNCVSIALLGVLLGAAALPRGAFAFEIALPRESNITVLIKGTPTLLAKPCYIYI 89
QY 61 SKRHITMLSIKSGERIVFTFSCSPENHFIYIEIKNIDCMGSPGPEVOLOPSTSLPT 120
DB 90 SKRHITMLSIKSGERIVFTFSCSPENHFIYIEIKNIDCMGSPGPEVOLOPSTSLPT 149
QY 121 LNRFTIWDVKAHKSIGLEIQFSIPRLRQIGPGBSCPDVYTHISGRIDATVVRIGTFCSN 180
DB 150 LNRFTIWDVKAHKSIGLEIQFSIPRLRQIGPGBSCPDVYTHISGRIDATVVRIGTFCSN 209
QY 181 GYVSRIKMGEGVMAHLMPFHRNVSQFSIANRSSIKLCTIESFEEGSGATLMSANY 240
DB 210 GYVSRIKMGEGVMAHLMPFHRNVSQFSIANRSSIKLCTIESFEEGSGATLMSANY 269
QY 241 PEGFPEDELMTQFVPAHLRASVSFLNRLSNCEKREVEYIIGSTTNPVEFKLEDK 300
DB 270 PEGFPEDELMTQFVPAHLRASVSFLNRLSNCEKREVEYIIGSTTNPVEFKLEDK 329
QY 301 QPNNMAGNFNLSLQGGDDDAQSPGILRLQFOVLVQHPONESNKIYVVDLSNERAMSLTTE 360
DB 330 QPNNMAGNFNLSLQGGDDDAQSPGILRLQFOVLVQHPONESNKIYVVDLSNERAMSLTTE 389
QY 361 PRPVKQSRKRVPCGFCVLESRTCSNLTLTSGSKHKISFLCDDLTRLMNVK 413

DB 390 PRPVKQSRKRVPCGFCVLESRTCSNLTLTSGSKHKISFLCDDLTRLMNVK 442

RESULT 3

US-09-489-847-132

; Sequence 132, Application US/09489847

; Patent No. 6476195

; GENERAL INFORMATION:

; APPLICANT: Rosen et al

; TITLE OF INVENTION: 98 Human Secreted Proteins

; FILE REFERENCE: P2031P1

; CURRENT APPLICATION NUMBER: US/09/489,847

; CURRENT FILING DATE: 2000-01-24

; EARLIER APPLICATION NUMBER: PCT/US99/17130

; EARLIER FILING DATE: 1999-07-29

; EARLIER APPLICATION NUMBER: 60/094,657

; EARLIER FILING DATE: 1998-07-30

; EARLIER APPLICATION NUMBER: 60/095,486

; EARLIER FILING DATE: 1998-08-05

; EARLIER APPLICATION NUMBER: 60/096,319

; EARLIER FILING DATE: 1998-08-12

; EARLIER APPLICATION NUMBER: 60/095,454

; EARLIER FILING DATE: 1998-08-06

; EARLIER APPLICATION NUMBER: 60/095,455

; EARLIER FILING DATE: 1998-08-06

; NUMBER OF SEQ ID NOS: 376

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 132

; LENGTH: 709

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (189)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (275)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (414)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (438)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (641)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (643)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (696)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (697)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; US-09-489-847-132

Query Match 25.2%; Score 211; DB 4; Length 709;

Best Local Similarity 99.5%; Pred. No. 2.2e-203;

Matches 411; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAGLNCVSIALLGVLLGAAALPRGAFAFEIALPRESNITVLIKGTPTLLAKPCYIYI 60
DB 1 MAGLNCVSIALLGVLLGAAALPRGAFAFEIALPRESNITVLIKGTPTLLAKPCYIYI 60

APPLICANT: Quigley, James P.
APPLICANT: Seandel, Marco
TITLE OF INVENTION: MONOCLONAL ANTIBODIES THAT RECOGNIZE ANTIGENS
TITLE OF INVENTION: ASSOCIATED WITH TUMOR METASTASIS
FILE REFERENCE: SUNY
CURRENT APPLICATION NUMBER: US/09/499,781
CURRENT FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: 09/333,599
PRIOR FILING DATE: 1999-06-15
NUMBER OF SEQ ID NOS: 5
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 3
LENGTH: 19
TYPE: PRT
ORGANISM: synthetic construct
FEATURE:
NAME/KEY: UNSURE
LOCATION: (15)
OTHER INFORMATION: Xaa is Gln or Ile.
NAME/KEY: UNSURE
LOCATION: (17)
US-09-499-781-3

Query Match 1.0%; Score 8; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 FEIAPRES 38
DB 1 FEIAPRES 9

RESULT 8
US-09-369-247-79
Sequence 79, Application US/09369247
Patent No. 6569992
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 44 Human Secreted Proteins
FILE REFERENCE: P2024P1
CURRENT APPLICATION NUMBER: US/09/369,247
CURRENT FILING DATE: 1999-08-05
EARLIER APPLICATION NUMBER: 60/074,118
EARLIER FILING DATE: 1998-02-09
EARLIER APPLICATION NUMBER: 60/074,157
EARLIER FILING DATE: 1998-02-09
EARLIER APPLICATION NUMBER: 60/074,137
EARLIER FILING DATE: 1998-02-09
EARLIER APPLICATION NUMBER: 60/074,341
EARLIER FILING DATE: 1998-02-09
EARLIER APPLICATION NUMBER: 60/074,141
EARLIER FILING DATE: 1998-02-09
NUMBER OF SEQ ID NOS: 172
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 79
LENGTH: 71
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (71)
OTHER INFORMATION: Xaa equals stop translation
US-09-369-247-79

Query Match 1.0%; Score 8; DB 4; Length 71;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AGINCGVS 9
DB 41 AGINCGVS 48

RESULT 9
US-09-252-991A-32450
Sequence 32450, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 32450
LENGTH: 156
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32450

Query Match 1.0%; Score 8; DB 4; Length 156;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GAARLPRG 26
DB 32 GAARLPRG 39

RESULT 10
US-09-599-360B-104
Sequence 104, Application US/09599360B
Patent No. 6548633
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Bouquelerec, L.
TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
FILE REFERENCE: GENSET.050CP3
CURRENT APPLICATION NUMBER: US/09/599,360B
CURRENT FILING DATE: 2000-06-21
PRIOR APPLICATION NUMBER: 60/113,686
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/141,032
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 09/469,099
PRIOR FILING DATE: 1999-12-21
NUMBER OF SEQ ID NOS: 123
SOFTWARE: Patent. pm
SEQ ID NO 104
LENGTH: 221
TYPE: PRT
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -28...-1
US-09-599-360B-104

Query Match 1.0%; Score 8; DB 4; Length 221;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LUGVLLG 19
DB 104 LUGVLLG 111

RESULT 11
US-09-252-991A-17865
Sequence 17865, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 17865
LENGTH: 703
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17865

Query Match 1.0%; Score 8; DB 4; Length 703;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 ALLGVLL 18
DB 386 ALLGVLL 393

RESULT 12
US-09-082-358B-43
Sequence 43, Application US/09082358B
Patent No. 6469153
GENERAL INFORMATION:
APPLICANT: Goff, Stephen P.
TITLE OF INVENTION: EIP-1, EIP-3 GENES, ENVELOPE-INTERACTING PROTEINS,
FILE REFERENCE: 0575/54804
CURRENT APPLICATION NUMBER: US/09/082,358B
CURRENT FILING DATE: 1998-05-20
NUMBER OF SEQ ID NOS: 106
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 43
LENGTH: 57
TYPE: PRT
ORGANISM: murine
US-09-082-358B-43

Query Match 0.8%; Score 7; DB 4; Length 57;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 692 KKKKKKT 698
DB 50 KKKKKKT 56

RESULT 13
US-09-205-258-381
Sequence 381, Application US/09205258
Patent No. 6525174
GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: P200781
CURRENT APPLICATION NUMBER: US/09/205,258
CURRENT FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: PCT/US98/11422
EARLIER FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: 60/048,885
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,375
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,881

EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,880
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,896
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,020
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,876
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,884
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,894
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,971
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,882
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,899
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30

```
/ NUMBER OF SEQ ID NOS: 1227
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 381
/ LENGTH: 66
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (14)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (62)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-205-258-381

Query Match
Best Local Similarity 100.0%; Score 7; DB 4; Length 66;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 692 KKKKKKT 698
Db 49 KKKKKKT 55

RESULT 14
US-09-732-210-341
/ Sequence 341, Application US/09732210
/ Patent No. 6573361
/ GENERAL INFORMATION:
/ APPLICANT: Bunkers, Greg J.
/ APPLICANT: Liang, Jihong
/ APPLICANT: Mitcank, Cindy A.
/ APPLICANT: Seale, Jeffrey W.
/ APPLICANT: Wu, Yonnie S.
/ TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
/ FILE REFERENCE: 38-21(15036)B
/ CURRENT APPLICATION NUMBER: US/09/732,210
/ PRIOR FILING DATE: 2000-12-07
/ PRIOR APPLICATION NUMBER: US 60/169,513
/ PRIOR FILING DATE: 1999-12-07
/ PRIOR APPLICATION NUMBER: US 60/169,340
/ NUMBER OF SEQ ID NOS: 1753
/ SEQ ID NO 341
/ LENGTH: 78
/ TYPE: PRT
/ ORGANISM: Dictyostelium discoideum
US-09-732-210-341

Query Match
Best Local Similarity 100.0%; Score 7; DB 4; Length 78;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 692 KKKKKKT 698
Db 3 KKKKKKT 9

RESULT 15
US-08-946-329A-94
/ Sequence 94, Application US/08946329A
/ Patent No. 6057091
/ GENERAL INFORMATION:
/ APPLICANT: Beachy, Philip A.
/ APPLICANT: Porter, Jeffrey A.
/ TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
/ NUMBER OF SEQUENCES: 109
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Fish & Richardson P.C.
/ STREET: 4225 Executive Square, Suite 1400
/ CITY: La Jolla
/ STATE: CA
```

```
/ COUNTRY: USA
/ ZIP: 92037
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: Windows 95
/ SOFTWARE: FASTSEQ for Windows Version 2.0b
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/946,329A
/ FILING DATE: 07-OCT-1997
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 60/061,323
/ FILING DATE: 07-OCT-1996
/ APPLICATION NUMBER: 08/729,743
/ FILING DATE: 10-JUL-1996
/ APPLICATION NUMBER: 08/567,357
/ FILING DATE: 04-DEC-1995
/ APPLICATION NUMBER: 08/349,498
/ FILING DATE: 02-DEC-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Haile, Lisa A.
/ REGISTRATION NUMBER: 38,347
/ REFERENCE/DOCKET NUMBER: 07265/140001
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 619/678-5070
/ TELEFAX: 619/678-5099
/ INFORMATION FOR SEQ ID NO: 94:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 97 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
US-08-946-329A-94

Query Match
Best Local Similarity 100.0%; Score 7; DB 3; Length 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 615 AEEIFSL 621
Db 17 AEEIFSL 23
```

Search completed: February 20, 2004, 19:51:00
Job time : 33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2004, 19:46:38 ; Search time 100 Seconds
(without alignments)
1750.436 Million cell updates/sec

Title: US-09-899-569A-4

Perfect score: 836
Sequence: 1 MAGLNCGVSIALLGVLLGA.....SSKQTDIPULTQEPMEPAE 836

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 801455 seqs, 209382283 residues

Word size : 0

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	836	100.0	836	US-09-899-569A-4	Sequence 4, Appl1
2	650	77.8	749	US-09-899-569A-2	Sequence 2, Appl1
3	341	40.8	343	US-10-137-870-162	Sequence 162, App
4	341	40.8	343	US-10-140-018-162	Sequence 162, App
5	341	40.8	343	US-10-140-021-162	Sequence 162, App
6	341	40.8	343	US-10-140-274-162	Sequence 162, App
7	341	40.8	343	US-10-140-471-162	Sequence 162, App
8	341	40.8	343	US-10-140-807-162	Sequence 162, App
9	341	40.8	343	US-10-140-922-162	Sequence 162, App
10	341	40.8	343	US-10-140-924-162	Sequence 162, App
11	341	40.8	343	US-10-140-926-162	Sequence 162, App
12	341	40.8	343	US-10-141-698-162	Sequence 162, App
13	341	40.8	343	US-10-141-702-162	Sequence 162, App
14	341	40.8	343	US-10-141-704-162	Sequence 162, App
15	341	40.8	343	US-10-142-421-162	Sequence 162, App

16	341	40.8	343	US-10-142-432-162	Sequence 162, App
17	341	40.8	343	US-10-142-767-162	Sequence 162, App
18	341	40.8	343	US-10-143-033-162	Sequence 162, App
19	341	40.8	343	US-10-144-994-162	Sequence 162, App
20	341	40.8	343	US-10-145-628-162	Sequence 162, App
21	341	40.8	343	US-10-145-631-162	Sequence 162, App
22	341	40.8	343	US-10-145-633-162	Sequence 162, App
23	341	40.8	343	US-10-145-746-162	Sequence 162, App
24	341	40.8	343	US-10-145-748-162	Sequence 162, App
25	341	40.8	343	US-10-145-823-162	Sequence 162, App
26	341	40.8	343	US-10-145-826-162	Sequence 162, App
27	341	40.8	343	US-10-145-870-162	Sequence 162, App
28	341	40.8	343	US-10-145-876-162	Sequence 162, App
29	341	40.8	343	US-10-145-959-162	Sequence 162, App
30	341	40.8	343	US-10-146-724-162	Sequence 162, App
31	341	40.8	343	US-10-146-725-162	Sequence 162, App
32	341	40.8	343	US-10-146-795-162	Sequence 162, App
33	341	40.8	343	US-10-147-495-162	Sequence 162, App
34	341	40.8	343	US-10-147-501-162	Sequence 162, App
35	341	40.8	343	US-10-147-504-162	Sequence 162, App
36	341	40.8	343	US-10-147-506-162	Sequence 162, App
37	341	40.8	343	US-10-147-509-162	Sequence 162, App
38	341	40.8	343	US-10-147-510-162	Sequence 162, App
39	341	40.8	343	US-10-147-511-162	Sequence 162, App
40	341	40.8	343	US-10-147-529-162	Sequence 162, App
41	341	40.8	343	US-10-152-397-162	Sequence 162, App
42	341	40.8	343	US-10-153-586-162	Sequence 162, App
43	341	40.8	343	US-10-158-793-162	Sequence 162, App
44	341	40.8	343	US-10-158-786-162	Sequence 162, App
45	341	40.8	343	US-10-140-019-162	Sequence 162, App

ALIGNMENTS

RESULT 1
US-09-899-569A-4
; Sequence 4, Application US/0989569A
; Patent No. US20020142003A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020142003Albert Schweitzer
; APPLICANT: Marwa Scherl-Mostageer
; APPLICANT: Wolfgang Sommergruber
; APPLICANT: Roger Absbeher
; TITLE OF INVENTION: Tumorsozioleites Antigen (B345)
; FILE REFERENCE: 0652.2280001
; CURRENT APPLICATION NUMBER: US/09/899,569A
; CURRENT FILING DATE: 2001-07-06
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: DE 100 33 080.0
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 60/243,158
; PRIOR FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: US 60/297,747
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-899-569A-4
Query Match 100.0%; Score 836; DB 10; Length 836;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 836; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLGAARLPAGAFETALPRESNITVLTGPTTLAKKCYVI 60
DB 1 MAGLNCGVSIALLGVLLGAARLPAGAFETALPRESNITVLTGPTTLAKKCYVI 60
QY 61 SKRHITMLSIKSGRIIVTFSCQSPENHFVIEIQNIDMGSPCPFGVQLQPSLPT 120

```

Db      61 SKRHITMLSIKSGERIVFTFSCSPENHVFIEIQKNIDCMGCPGEGVQLOPSTSLLP 120
Qy      121 LNRTFIMDKAHKSIIGLELOFSIRLROIGPGECPDGVTHSISGRIDATVVRIGTFCSN 180
Db      121 LNRTFIMDKAHKSIIGLELOFSIRLROIGPGECPDGVTHSISGRIDATVVRIGTFCSN 180
Qy      181 GIVSRIRKMOEGVMKALHLPMPHPRNVSGFSIANRSSIKRLCTIESEVEGEGSATLMSANY 240
Db      181 GIVSRIRKMOEGVMKALHLPMPHPRNVSGFSIANRSSIKRLCTIESEVEGEGSATLMSANY 240
Qy      241 PGGFPEDELTMOQFVPAHLRASVSFLNLSNCEKKEEVEYIIGSTINPEVFLDEK 300
Db      241 PGGFPEDELTMOQFVPAHLRASVSFLNLSNCEKKEEVEYIIGSTINPEVFLDEK 300
Qy      301 QPGNMAGNNTLSQGCDDAOSPGILRLOFVLOHPONESNKIYVDLSNERAMSLTIE 360
Db      301 QPGNMAGNNTLSQGCDDAOSPGILRLOFVLOHPONESNKIYVDLSNERAMSLTIE 360
Qy      361 PRPVKSRKFPVCGFVLESRTCSNLTJSGSKHISFLCDDLRLMNMVEXTISCTDH 420
Db      361 PRPVKSRKFPVCGFVLESRTCSNLTJSGSKHISFLCDDLRLMNMVEXTISCTDH 420
Qy      421 RVCQRSSYSLOVPSDILHLPVELHDFSXKLLVPKRLSLVPAQKLOQTHEKPCNTSF 480
Db      421 RVCQRSSYSLOVPSDILHLPVELHDFSXKLLVPKRLSLVPAQKLOQTHEKPCNTSF 480
Qy      481 SYLVASAISSODLYFSGFCPGGSIKOIOVKONISVLTFTFAPSFOEASROGLTVSFIFY 540
Db      481 SYLVASAISSODLYFSGFCPGGSIKOIOVKONISVLTFTFAPSFOEASROGLTVSFIFY 540
Qy      541 FKEEGVFTVPTDKSKYVLRTPNMDRGLPSLTSVSNISVPRDQVACLTFPKERSGVCC 600
Db      541 FKEEGVFTVPTDKSKYVLRTPNMDRGLPSLTSVSNISVPRDQVACLTFPKERSGVCC 600
Qy      601 TGRAPMIIQEBORTRAEISLDEVDLPKPSFHHHSFWNINISNCSPTSGKQDLDFSVTLT 660
Db      601 TGRAPMIIQEBORTRAEISLDEVDLPKPSFHHHSFWNINISNCSPTSGKQDLDFSVTLT 660
Qy      661 PRIVDITVLILAAVGGVLLLSALGLIICVKKKKKTKNKPAGVINGNINTEMPOR 720
Db      661 PRIVDITVLILAAVGGVLLLSALGLIICVKKKKKTKNKPAGVINGNINTEMPOR 720
Qy      721 KFOGKGRKNDSHYAVIEDTMVYGLLQDSGSLQPEVDYTRPFQGTMGVCPSPPTIC 780
Db      721 KFOGKGRKNDSHYAVIEDTMVYGLLQDSGSLQPEVDYTRPFQGTMGVCPSPPTIC 780
Qy      781 SRAPTAKLATEEPSPSPSESEPTTFSHPNNGDVSSKOTDIPLINTQEPMEPAR 836
Db      781 SRAPTAKLATEEPSPSPSESEPTTFSHPNNGDVSSKOTDIPLINTQEPMEPAR 836

```

RESULT 2
US-09-899-569a-2

```

; Sequence 2, Application US/09899569A
; Patent No. US20020142003A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020142003Albert Schweifer
; APPLICANT: Marwa Scherl-Mostageer
; APPLICANT: Wolfgang Sommergruber
; APPLICANT: Roger Absener
; TITLE OF INVENTION: Tumoraesoziiertes Antigen (B345)
; FILE REFERENCE: 0652.2280001
; CURRENT APPLICATION NUMBER: US/09/899,569A
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: DE 100 33 080.0
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: DE 101 19 294.0
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 60/243,158
; PRIOR FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: US 60/297,747
; PRIOR FILING DATE: 2001-06-14

```

```

; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 749
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-899-569a-2

```

Query Match 77.8%; Score 650; DB 10; Length 749;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 650; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      67 MLIISGERIVFTFSCSPENHVFIEIQKNIDCMGCPGEGVQLOPSTSLLP 126
Db      1 MLIISGERIVFTFSCSPENHVFIEIQKNIDCMGCPGEGVQLOPSTSLLP 126
Qy      127 WDVKAHKSIGLELOFSIRLROIGPGECPDGVTHSISGRIDATVVRIGTFCSN 186
Db      61 WDVKAHKSIGLELOFSIRLROIGPGECPDGVTHSISGRIDATVVRIGTFCSN 186
Qy      187 KMOEGVMKALHLPMPHPRNVSGFSIANRSSIKRLCTIESEVEGEGSATLMSANYEGPPE 246
Db      121 KMOEGVMKALHLPMPHPRNVSGFSIANRSSIKRLCTIESEVEGEGSATLMSANYEGPPE 180
Qy      247 DELMTQFVPAHLRASVSFLNLSNCEKKEEVEYIIGSTINPEVFLKEDKOPGMA 306
Db      181 DELMTQFVPAHLRASVSFLNLSNCEKKEEVEYIIGSTINPEVFLKEDKOPGMA 240
Qy      307 GNFNLSQGCDDAOSPGILRLOFVLOHPONESNKIYVDLSNERAMSLTIEPRVQ 366
Db      241 GNFNLSQGCDDAOSPGILRLOFVLOHPONESNKIYVDLSNERAMSLTIEPRVQ 300
Qy      367 SRKFVPGFVCLSESTCSNLTJSGSKHISFLCDDLRLMNMVEXTISCTDHRVCOR 426
Db      301 SRKFVPGFVCLSESTCSNLTJSGSKHISFLCDDLRLMNMVEXTISCTDHRVCOR 360
Qy      427 SYSLQVPSDILHLPVELHDFSXKLLVPKRLSLVPAQKLOQTHEKPCNTSFYLVAS 486
Db      361 SYSLQVPSDILHLPVELHDFSXKLLVPKRLSLVPAQKLOQTHEKPCNTSFYLVAS 420
Qy      487 AIPSDLYFSGFCPGGSIKOIOVKONISVLTFTFAPSFOEASROGLTVSFIFYKEEGV 546
Db      421 AIPSDLYFSGFCPGGSIKOIOVKONISVLTFTFAPSFOEASROGLTVSFIFYKEEGV 480
Qy      547 FTVPTDKSKYVLRTPNMDRGLPSLTSVSNISVPRDQVACLTFPKERSGVCCQGRAPM 606
Db      481 FTVPTDKSKYVLRTPNMDRGLPSLTSVSNISVPRDQVACLTFPKERSGVCCQGRAPM 540
Qy      607 IIOEBORTRAEISLDEVDLPKPSFHHHSFWNINISNCSPTSGKQDLDFSVTLTPTV 666
Db      541 IIOEBORTRAEISLDEVDLPKPSFHHHSFWNINISNCSPTSGKQDLDFSVTLTPTV 600
Qy      667 TVILIAAVGGVLLLSALGLIICVKKKKKTKNKPAGVINGNINTEMP 716
Db      601 TVILIAAVGGVLLLSALGLIICVKKKKKTKNKPAGVINGNINTEMP 650

```

RESULT 3

```

US-10-137-870-162
; Sequence 162, Application US/10137870
; Publication No. US20030138883A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Geo, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven

```

APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C155
CURRENT APPLICATION NUMBER: US/10/137,870
CURRENT FILING DATE: 2002-05-03
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 162
LENGTH: 343
TYPE: PRT
ORGANISM: Homo Sapien
US-10-137-870-162

Query Match 40.8%; Score 341; DB 12; Length 343;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNGVSIALGVLLGAARLPGAAFEIALPRESNITVLKLGPTTLAKPCYIV 60
DB 1 MAGLNGVSIALGVLLGAARLPGAAFEIALPRESNITVLKLGPTTLAKPCYIV 60
QY 61 SKRHITMISIGSERIVTFPSCQSPENHFVIEIQKIDMSGPCPGEVQLPSTSLPT 120
DB 61 SKRHITMISIGSERIVTFPSCQSPENHFVIEIQKIDMSGPCPGEVQLPSTSLPT 120
QY 121 LNRTFMVKAHKSIGLEQFSIPRLROIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRTFMVKAHKSIGLEQFSIPRLROIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKQEGVKMALHPWFHPRVSGFSIANRSIRKLCIESVFEGEGSATLMSANY 240
DB 181 GTVSRIKQEGVKMALHPWFHPRVSGFSIANRSIRKLCIESVFEGEGSATLMSANY 240
QY 241 PEGPEDELMTQFVVPALRASVSFLNPNLNCERKEERYIIPGSTTNEVFKLBDK 300
DB 241 PEGPEDELMTQFVVPALRASVSFLNPNLNCERKEERYIIPGSTTNEVFKLBDK 300
QY 301 QPGNMGNFNLSLQCDQDQSPGILRLQFVLVQHPONES 341
DB 301 QPGNMGNFNLSLQCDQDQSPGILRLQFVLVQHPONES 341

RESULT 4

US-10-140-018-162
Sequence 162, Application US/10140018
Publication No. US20030138865A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C158

CURRENT APPLICATION NUMBER: US/10/140,018
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 162
LENGTH: 343
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-018-162

Query Match 40.8%; Score 341; DB 12; Length 343;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNGVSIALGVLLGAARLPGAAFEIALPRESNITVLKLGPTTLAKPCYIV 60
DB 1 MAGLNGVSIALGVLLGAARLPGAAFEIALPRESNITVLKLGPTTLAKPCYIV 60
QY 61 SKRHITMISIGSERIVTFPSCQSPENHFVIEIQKIDMSGPCPGEVQLPSTSLPT 120
DB 61 SKRHITMISIGSERIVTFPSCQSPENHFVIEIQKIDMSGPCPGEVQLPSTSLPT 120
QY 121 LNRTFMVKAHKSIGLEQFSIPRLROIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRTFMVKAHKSIGLEQFSIPRLROIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKQEGVKMALHPWFHPRVSGFSIANRSIRKLCIESVFEGEGSATLMSANY 240
DB 181 GTVSRIKQEGVKMALHPWFHPRVSGFSIANRSIRKLCIESVFEGEGSATLMSANY 240
QY 241 PEGPEDELMTQFVVPALRASVSFLNPNLNCERKEERYIIPGSTTNEVFKLBDK 300
DB 241 PEGPEDELMTQFVVPALRASVSFLNPNLNCERKEERYIIPGSTTNEVFKLBDK 300
QY 301 QPGNMGNFNLSLQCDQDQSPGILRLQFVLVQHPONES 341
DB 301 QPGNMGNFNLSLQCDQDQSPGILRLQFVLVQHPONES 341

RESULT 5

US-10-140-021-162
Sequence 162, Application US/10140021
Publication No. US2003013886A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C167
CURRENT APPLICATION NUMBER: US/10/140,021
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 162
LENGTH: 343
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-021-162

Query Match 40.8%; Score 341; DB 12; Length 343;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNGVSIALGVLGAAARLPAGAFAETALPRESNITVLKGTPTLAKPCYIYI 60
DB 1 MAGLNGVSIALGVLGAAARLPAGAFAETALPRESNITVLKGTPTLAKPCYIYI 60
QY 61 SKRHTMTLSKGERIVTFSCQSPENHFVIEIQKIDCMGSPCPFGVQLQPSSTLLPT 120
DB 61 SKRHTMTLSKGERIVTFSCQSPENHFVIEIQKIDCMGSPCPFGVQLQPSSTLLPT 120
QY 121 LNRFTIMVKAHKSIGLELQFSIPRLROIIPGESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIMVKAHKSIGLELQFSIPRLROIIPGESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKMOEGVKMALHPWFHPRNVSGFSIANRSSIKRLCTIESVFEGSGSATLMSANY 240
DB 181 GTVSRIKMOEGVKMALHPWFHPRNVSGFSIANRSSIKRLCTIESVFEGSGSATLMSANY 240
QY 241 PEGPPEDELMTWQFVVPALHRAVSFLNPNLSNCRKKEERYIIPGSTTNEVFKLBDK 300
DB 241 PEGPPEDELMTWQFVVPALHRAVSFLNPNLSNCRKKEERYIIPGSTTNEVFKLBDK 300
QY 301 QPGNAGNFNLSLQCGDDQDQASPGILRLQFOVLVQHPONES 341
DB 301 QPGNAGNFNLSLQCGDDQDQASPGILRLQFOVLVQHPONES 341

RESULT 6
US-10-140-274-162
; Sequence 162, Application US/10140274
; Publication No. US20030143674A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Goddard, Paul J.
APPLICANT: Gueney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C161
CURRENT APPLICATION NUMBER: US/10/140,274
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 162
LENGTH: 343
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-274-162

Query Match 40.8%; Score 341; DB 12; Length 343;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNGVSIALGVLGAAARLPAGAFAETALPRESNITVLKGTPTLAKPCYIYI 60
DB 1 MAGLNGVSIALGVLGAAARLPAGAFAETALPRESNITVLKGTPTLAKPCYIYI 60

QY 61 SKRHTMTLSKGERIVTFSCQSPENHFVIEIQKIDCMGSPCPFGVQLQPSSTLLPT 120
DB 61 SKRHTMTLSKGERIVTFSCQSPENHFVIEIQKIDCMGSPCPFGVQLQPSSTLLPT 120
QY 121 LNRFTIMVKAHKSIGLELQFSIPRLROIIPGESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIMVKAHKSIGLELQFSIPRLROIIPGESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKMOEGVKMALHPWFHPRNVSGFSIANRSSIKRLCTIESVFEGSGSATLMSANY 240
DB 181 GTVSRIKMOEGVKMALHPWFHPRNVSGFSIANRSSIKRLCTIESVFEGSGSATLMSANY 240
QY 241 PEGPPEDELMTWQFVVPALHRAVSFLNPNLSNCRKKEERYIIPGSTTNEVFKLBDK 300
DB 241 PEGPPEDELMTWQFVVPALHRAVSFLNPNLSNCRKKEERYIIPGSTTNEVFKLBDK 300
QY 301 QPGNAGNFNLSLQCGDDQDQASPGILRLQFOVLVQHPONES 341
DB 301 QPGNAGNFNLSLQCGDDQDQASPGILRLQFOVLVQHPONES 341

RESULT 7
US-10-140-471-162
; Sequence 162, Application US/10140471
; Publication No. US2003013887A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Gueney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C163
CURRENT APPLICATION NUMBER: US/10/140,471
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 162
LENGTH: 343
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-471-162

Query Match 40.8%; Score 341; DB 12; Length 343;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNGVSIALGVLGAAARLPAGAFAETALPRESNITVLKGTPTLAKPCYIYI 60
DB 1 MAGLNGVSIALGVLGAAARLPAGAFAETALPRESNITVLKGTPTLAKPCYIYI 60
QY 61 SKRHTMTLSKGERIVTFSCQSPENHFVIEIQKIDCMGSPCPFGVQLQPSSTLLPT 120
DB 61 SKRHTMTLSKGERIVTFSCQSPENHFVIEIQKIDCMGSPCPFGVQLQPSSTLLPT 120
QY 121 LNRFTIMVKAHKSIGLELQFSIPRLROIIPGESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIMVKAHKSIGLELQFSIPRLROIIPGESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKMOEGVKMALHPWFHPRNVSGFSIANRSSIKRLCTIESVFEGSGSATLMSANY 240

Db 181 GTVSRKIQEGVKALHLPWFHPRNVSGFSIANSSIRLCIIISVFEGESATLMSANY 240
QY 241 PEGPPEDELMTWQFVPAHLRAVSFLNFNLSNCRKEERYEYIPGSTTNPEVFKLEDK 300-
Db 241 PEGPPEDELMTWQFVPAHLRAVSFLNFNLSNCRKEERYEYIPGSTTNPEVFKLEDK 300
QY 301 QPGMAGNFNLSLOGCDQDQASPGILRLQFOVLVQHONES 341
Db 301 QPGMAGNFNLSLOGCDQDQASPGILRLQFOVLVQHONES 341
RESULT 8
US-10-140-807-162
; Sequence 162, Application US/10140807
; Publication No. US20030134354A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C174
; CURRENT APPLICATION NUMBER: US/10/140,807
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-807-162
Query Match 40.8%; Score 341; DB 12; Length 343;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAGNCGVSIALLVLLGAARLPFGAFAFEIALPRESNTIVLTKGTPTLLAKPCYIVI 60
Db 1 MAGNCGVSIALLVLLGAARLPFGAFAFEIALPRESNTIVLTKGTPTLLAKPCYIVI 60
QY 61 SKRHITMISIGSERIVFTFSCQSPENHFVIEIQKIDCMGSPCPFGVQLQPSISLPT 120
Db 61 SKRHITMISIGSERIVFTFSCQSPENHFVIEIQKIDCMGSPCPFGVQLQPSISLPT 120
QY 121 LNRFTFMDVKAKHSIGLEQFSIPRLRQIGGESCPDGVTHSIGRIDATVVRIGTFPCSN 180
Db 121 LNRFTFMDVKAKHSIGLEQFSIPRLRQIGGESCPDGVTHSIGRIDATVVRIGTFPCSN 180
QY 181 GTVSRKIQEGVKALHLPWFHPRNVSGFSIANSSIRLCIIISVFEGESATLMSANY 240
Db 181 GTVSRKIQEGVKALHLPWFHPRNVSGFSIANSSIRLCIIISVFEGESATLMSANY 240
QY 241 PEGPPEDELMTWQFVPAHLRAVSFLNFNLSNCRKEERYEYIPGSTTNPEVFKLEDK 300
Db 241 PEGPPEDELMTWQFVPAHLRAVSFLNFNLSNCRKEERYEYIPGSTTNPEVFKLEDK 300
QY 301 QPGMAGNFNLSLOGCDQDQASPGILRLQFOVLVQHONES 341
Db 301 QPGMAGNFNLSLOGCDQDQASPGILRLQFOVLVQHONES 341

Db 301 QPGMAGNFNLSLOGCDQDQASPGILRLQFOVLVQHONES 341
RESULT 9
US-10-140-922-162
; Sequence 162, Application US/10140922
; Publication No. US20030138889A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C179
; CURRENT APPLICATION NUMBER: US/10/140,922
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-922-162
Query Match 40.8%; Score 341; DB 12; Length 343;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAGNCGVSIALLVLLGAARLPFGAFAFEIALPRESNTIVLTKGTPTLLAKPCYIVI 60
Db 1 MAGNCGVSIALLVLLGAARLPFGAFAFEIALPRESNTIVLTKGTPTLLAKPCYIVI 60
QY 61 SKRHITMISIGSERIVFTFSCQSPENHFVIEIQKIDCMGSPCPFGVQLQPSISLPT 120
Db 61 SKRHITMISIGSERIVFTFSCQSPENHFVIEIQKIDCMGSPCPFGVQLQPSISLPT 120
QY 121 LNRFTFMDVKAKHSIGLEQFSIPRLRQIGGESCPDGVTHSIGRIDATVVRIGTFPCSN 180
Db 121 LNRFTFMDVKAKHSIGLEQFSIPRLRQIGGESCPDGVTHSIGRIDATVVRIGTFPCSN 180
QY 181 GTVSRKIQEGVKALHLPWFHPRNVSGFSIANSSIRLCIIISVFEGESATLMSANY 240
Db 181 GTVSRKIQEGVKALHLPWFHPRNVSGFSIANSSIRLCIIISVFEGESATLMSANY 240
QY 241 PEGPPEDELMTWQFVPAHLRAVSFLNFNLSNCRKEERYEYIPGSTTNPEVFKLEDK 300
Db 241 PEGPPEDELMTWQFVPAHLRAVSFLNFNLSNCRKEERYEYIPGSTTNPEVFKLEDK 300
QY 301 QPGMAGNFNLSLOGCDQDQASPGILRLQFOVLVQHONES 341
Db 301 QPGMAGNFNLSLOGCDQDQASPGILRLQFOVLVQHONES 341
RESULT 10
US-10-140-924-162
; Sequence 162, Application US/10140924
; Publication No. US20030134355A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.

```

; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C17
; CURRENT APPLICATION NUMBER: US/10/140,924
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-924-162

```

```

Query Match      40.8%; Score 341; DB 12; Length 343;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MAGLNCGVSIALLGVLLGAARLPRGAFAFEIALPRESNITVLIKGTPTLLAKCYIYI 60
DB 1 MAGLNCGVSIALLGVLLGAARLPRGAFAFEIALPRESNITVLIKGTPTLLAKCYIYI 60
QY 61 SKRHITMTLSIKSGERIVTFSCQSPENHFVIEIQKNIDCMGSPCPFGVEVOLQPSISLPT 120
DB 61 SKRHITMTLSIKSGERIVTFSCQSPENHFVIEIQKNIDCMGSPCPFGVEVOLQPSISLPT 120
QY 121 LNRTFIWVKAHKSIGLEQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRTFIWVKAHKSIGLEQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIRKQEGVKALHPMFHPRVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 240
DB 181 GTVSRIRKQEGVKALHPMFHPRVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 240
QY 241 PEGPPEDELMTWQFVVPALRLASVSFLNPNLSNCRKEERYEYIIGSTTNPVEFKLEDK 300
DB 241 PEGPPEDELMTWQFVVPALRLASVSFLNPNLSNCRKEERYEYIIGSTTNPVEFKLEDK 300
QY 301 QPGNMAGNFNLSLQGCDDAOSPGILRLQFOVLVHPONES 341
DB 301 QPGNMAGNFNLSLQGCDDAOSPGILRLQFOVLVHPONES 341

```

RESULT 11

```

; Sequence 162, Application US/10140926
; Publication No. US20030134356A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.

```

```

; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C187
; CURRENT APPLICATION NUMBER: US/10/140,926
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-926-162

```

```

Query Match      40.8%; Score 341; DB 12; Length 343;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MAGLNCGVSIALLGVLLGAARLPRGAFAFEIALPRESNITVLIKGTPTLLAKCYIYI 60
DB 1 MAGLNCGVSIALLGVLLGAARLPRGAFAFEIALPRESNITVLIKGTPTLLAKCYIYI 60
QY 61 SKRHITMTLSIKSGERIVTFSCQSPENHFVIEIQKNIDCMGSPCPFGVEVOLQPSISLPT 120
DB 61 SKRHITMTLSIKSGERIVTFSCQSPENHFVIEIQKNIDCMGSPCPFGVEVOLQPSISLPT 120
QY 121 LNRTFIWVKAHKSIGLEQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRTFIWVKAHKSIGLEQFSIPRLRQIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIRKQEGVKALHPMFHPRVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 240
DB 181 GTVSRIRKQEGVKALHPMFHPRVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 240
QY 241 PEGPPEDELMTWQFVVPALRLASVSFLNPNLSNCRKEERYEYIIGSTTNPVEFKLEDK 300
DB 241 PEGPPEDELMTWQFVVPALRLASVSFLNPNLSNCRKEERYEYIIGSTTNPVEFKLEDK 300
QY 301 QPGNMAGNFNLSLQGCDDAOSPGILRLQFOVLVHPONES 341
DB 301 QPGNMAGNFNLSLQGCDDAOSPGILRLQFOVLVHPONES 341

```

RESULT 12

```

; Sequence 162, Application US/10141698
; Publication No. US20030134357A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

```

FILE REFERENCE: P3330R1C206
 CURRENT APPLICATION NUMBER: US/10/141.698
 CURRENT FILING DATE: 2002-05-08
 Prior Application removed - See Palm or File Wrapper
 NUMBER OF SEQ ID NOS: 550
 SEQ ID NO 162
 LENGTH: 343
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-10-141-698-162

Query Match 40.8%; Score 341; DB 12; Length 343;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALLVLLGAARLPFGAFAFIALPRESNITVLKLGPTLLAKRCYIVI 60
 DB 1 MAGNCGVSIALLVLLGAARLPFGAFAFIALPRESNITVLKLGPTLLAKRCYIVI 60
 QY 61 SKRHITMISIGSERIVTFSCQSPENHFVIEIQKNDMSGPCPFGEVOLQPSLSLPT 120
 DB 61 SKRHITMISIGSERIVTFSCQSPENHFVIEIQKNDMSGPCPFGEVOLQPSLSLPT 120
 QY 121 LNRFTIMVKAHKSIGLEQFSIPRLRQIGPESCPDGVTHSIGRIDATVVRIGTFPCSN 180
 DB 121 LNRFTIMVKAHKSIGLEQFSIPRLRQIGPESCPDGVTHSIGRIDATVVRIGTFPCSN 180
 QY 181 GTVSRIMKQEGVKMALHPWFHPRVNVSFSTANRSSIKRLCIISVFEGESATLMSANY 240
 DB 181 GTVSRIMKQEGVKMALHPWFHPRVNVSFSTANRSSIKRLCIISVFEGESATLMSANY 240
 QY 241 PEGFPEDELMTWQFVVPALHRAVSFLNPNLNSCERKEERYIIPGSTNPEVFKLEDK 300
 DB 241 PEGFPEDELMTWQFVVPALHRAVSFLNPNLNSCERKEERYIIPGSTNPEVFKLEDK 300
 QY 301 QPGNMGNFNLSLQCCDDAOSPGILRLQFQVLVQHPONES 341
 DB 301 QPGNMGNFNLSLQCCDDAOSPGILRLQFQVLVQHPONES 341

RESULT 13
 US-10-141-702-162
 Sequence 162, Application US/10141702
 Publication No. US20030134358A1
 GENERAL INFORMATION:
 APPLICANT: Baker, Kevin P.
 APPLICANT: Beresini, Maureen
 APPLICANT: DeForge, Laura
 APPLICANT: Desnoyers, Luc
 APPLICANT: Filvaroff, Ellen
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Gerritsen, Mary E.
 APPLICANT: Goddard, Audrey
 APPLICANT: Goddard, Paul J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Sherwood, Steven
 APPLICANT: Smith, Victoria
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Watanabe, Colin K
 APPLICANT: Wood, William
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 TITLE OF INVENTION: ACIDS ENCODING THE SAME
 FILE REFERENCE: P3330R1C208
 CURRENT APPLICATION NUMBER: US/10/141.702
 CURRENT FILING DATE: 2002-05-08
 Prior Application removed - See Palm or File Wrapper
 NUMBER OF SEQ ID NOS: 550
 SEQ ID NO 162
 LENGTH: 343
 TYPE: PRT
 ORGANISM: Homo Sapien

US-10-141-702-162

Query Match 40.8%; Score 341; DB 12; Length 343;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGNCGVSIALLVLLGAARLPFGAFAFIALPRESNITVLKLGPTLLAKRCYIVI 60
 DB 1 MAGNCGVSIALLVLLGAARLPFGAFAFIALPRESNITVLKLGPTLLAKRCYIVI 60
 QY 61 SKRHITMISIGSERIVTFSCQSPENHFVIEIQKNDMSGPCPFGEVOLQPSLSLPT 120
 DB 61 SKRHITMISIGSERIVTFSCQSPENHFVIEIQKNDMSGPCPFGEVOLQPSLSLPT 120
 QY 121 LNRFTIMVKAHKSIGLEQFSIPRLRQIGPESCPDGVTHSIGRIDATVVRIGTFPCSN 180
 DB 121 LNRFTIMVKAHKSIGLEQFSIPRLRQIGPESCPDGVTHSIGRIDATVVRIGTFPCSN 180
 QY 181 GTVSRIMKQEGVKMALHPWFHPRVNVSFSTANRSSIKRLCIISVFEGESATLMSANY 240
 DB 181 GTVSRIMKQEGVKMALHPWFHPRVNVSFSTANRSSIKRLCIISVFEGESATLMSANY 240
 QY 241 PEGFPEDELMTWQFVVPALHRAVSFLNPNLNSCERKEERYIIPGSTNPEVFKLEDK 300
 DB 241 PEGFPEDELMTWQFVVPALHRAVSFLNPNLNSCERKEERYIIPGSTNPEVFKLEDK 300
 QY 301 QPGNMGNFNLSLQCCDDAOSPGILRLQFQVLVQHPONES 341
 DB 301 QPGNMGNFNLSLQCCDDAOSPGILRLQFQVLVQHPONES 341

RESULT 14
 US-10-141-704-162
 Sequence 162, Application US/10141704
 Publication No. US20030134359A1
 GENERAL INFORMATION:
 APPLICANT: Baker, Kevin P.
 APPLICANT: Beresini, Maureen
 APPLICANT: DeForge, Laura
 APPLICANT: Desnoyers, Luc
 APPLICANT: Filvaroff, Ellen
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Gerritsen, Mary E.
 APPLICANT: Goddard, Audrey
 APPLICANT: Goddard, Paul J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Sherwood, Steven
 APPLICANT: Smith, Victoria
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Watanabe, Colin K
 APPLICANT: Wood, William
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 TITLE OF INVENTION: ACIDS ENCODING THE SAME
 FILE REFERENCE: P3330R1C209
 CURRENT APPLICATION NUMBER: US/10/141.704
 CURRENT FILING DATE: 2002-05-08
 Prior Application removed - See Palm or File Wrapper
 NUMBER OF SEQ ID NOS: 550
 SEQ ID NO 162
 LENGTH: 343
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-10-141-704-162
 Query Match 40.8%; Score 341; DB 12; Length 343;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```
QY 61 SRRHITMLSIKSGERIVFTFSCQSPENHFVIEIQXNIDCMGSPCPGVEVOLPSTSLPT 120
DB 61 SRRHITMLSIKSGERIVFTFSCQSPENHFVIEIQXNIDCMGSPCPGVEVOLPSTSLPT 120
QY 121 LNRFTIWDVKAHKSIGLEIQFSIPRLROIQPGSCPDGVTSHISGRIDATTVRIGTFPCSN 180
DB 121 LNRFTIWDVKAHKSIGLEIQFSIPRLROIQPGSCPDGVTSHISGRIDATTVRIGTFPCSN 180
QY 181 GTVSRIKMQEGVKMALHLPWFHPRNVSGFSIANRSSIKRLCTIESEVFESEGSATLMSANY 240
DB 181 GTVSRIKMQEGVKMALHLPWFHPRNVSGFSIANRSSIKRLCTIESEVFESEGSATLMSANY 240
QY 241 PEGFPEDELMTQFVVPALHRAVSFLNPNLSNCRKEERVEYYIPGSTTNPVEVKLEBK 300
DB 241 PEGFPEDELMTQFVVPALHRAVSFLNPNLSNCRKEERVEYYIPGSTTNPVEVKLEBK 300
QY 301 QPGNMAGNFNLSLOGCDQDQSPGILRLQFOVLVQHPONES 341
DB 301 QPGNMAGNFNLSLOGCDQDQSPGILRLQFOVLVQHPONES 341
```

RESULT 15

```
US-10-142-421-162
; Sequence 162, Application US/10142421
; Publication No. US20030134360A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Guiney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C218
; CURRENT APPLICATION NUMBER: US/10/142,421
; CURRENT FILING DATE: 2002-05-09
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 162
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-142-421-162
```

```
Query Match 40.8%; Score 341; DB 12; Length 343;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MAGLNCGVSIALLGVLLGAARLPKGAFAEIALPRESNITVLIKLGTPTLLAKPCYIV 60
DB 1 MAGLNCGVSIALLGVLLGAARLPKGAFAEIALPRESNITVLIKLGTPTLLAKPCYIV 60
QY 61 SRRHITMLSIKSGERIVFTFSCQSPENHFVIEIQXNIDCMGSPCPGVEVOLPSTSLPT 120
DB 61 SRRHITMLSIKSGERIVFTFSCQSPENHFVIEIQXNIDCMGSPCPGVEVOLPSTSLPT 120
QY 121 LNRFTIWDVKAHKSIGLEIQFSIPRLROIQPGSCPDGVTSHISGRIDATTVRIGTFPCSN 180
DB 121 LNRFTIWDVKAHKSIGLEIQFSIPRLROIQPGSCPDGVTSHISGRIDATTVRIGTFPCSN 180
```

```
QY 181 GTVSRIKMQEGVKMALHLPWFHPRNVSGFSIANRSSIKRLCTIESEVFESEGSATLMSANY 240
DB 181 GTVSRIKMQEGVKMALHLPWFHPRNVSGFSIANRSSIKRLCTIESEVFESEGSATLMSANY 240
QY 241 PEGFPEDELMTQFVVPALHRAVSFLNPNLSNCRKEERVEYYIPGSTTNPVEVKLEBK 300
DB 241 PEGFPEDELMTQFVVPALHRAVSFLNPNLSNCRKEERVEYYIPGSTTNPVEVKLEBK 300
QY 301 QPGNMAGNFNLSLOGCDQDQSPGILRLQFOVLVQHPONES 341
DB 301 QPGNMAGNFNLSLOGCDQDQSPGILRLQFOVLVQHPONES 341
```

Search completed: February 20, 2004, 19:52:51
Job time : 102 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using -bw model

Run on: February 20, 2004, 19:47:18 ; Search time 360 Seconds
(without alignments)
2113.034 Million cell updates/sec

US-09-899-569a-4

Title: 836
Perfect score: 1 MAGNCGSIALLCVLLIGA.....SSKDTDIPLANTORPEMPAE 836
Sequence:

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 5728757 seqs, 909918778 residues

Word size: 0

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: Pending_Patents_AA_Main:*

- 1: /cgn2_6/ptodaca/1/paa/US06_PCTUS_COMB.pep.*
- 2: /cgn2_6/ptodaca/1/paa/US06_COMB.pep.*
- 3: /cgn2_6/ptodaca/1/paa/US07_COMB.pep.*
- 4: /cgn2_6/ptodaca/1/paa/US08_COMB.pep.*
- 5: /cgn2_6/ptodaca/1/paa/US081_COMB.pep.*
- 6: /cgn2_6/ptodaca/1/paa/US082_COMB.pep.*
- 7: /cgn2_6/ptodaca/1/paa/US083_COMB.pep.*
- 8: /cgn2_6/ptodaca/1/paa/US084_COMB.pep.*
- 9: /cgn2_6/ptodaca/1/paa/US085_COMB.pep.*
- 10: /cgn2_6/ptodaca/1/paa/US086_COMB.pep.*
- 11: /cgn2_6/ptodaca/1/paa/US087_COMB.pep.*
- 12: /cgn2_6/ptodaca/1/paa/US088_COMB.pep.*
- 13: /cgn2_6/ptodaca/1/paa/US089_COMB.pep.*
- 14: /cgn2_6/ptodaca/1/paa/US090_COMB.pep.*
- 15: /cgn2_6/ptodaca/1/paa/US091_COMB.pep.*
- 16: /cgn2_6/ptodaca/1/paa/US092_COMB.pep.*
- 17: /cgn2_6/ptodaca/1/paa/US093_COMB.pep.*
- 18: /cgn2_6/ptodaca/1/paa/US094_COMB.pep.*
- 19: /cgn2_6/ptodaca/1/paa/US095_COMB.pep.*
- 20: /cgn2_6/ptodaca/1/paa/US096_COMB.pep.*
- 21: /cgn2_6/ptodaca/1/paa/US097_COMB.pep.*
- 22: /cgn2_6/ptodaca/1/paa/US097B_COMB.pep.*
- 23: /cgn2_6/ptodaca/1/paa/US098_COMB.pep.*
- 24: /cgn2_6/ptodaca/1/paa/US099_COMB.pep.*
- 25: /cgn2_6/ptodaca/1/paa/US099A_COMB.pep.*
- 26: /cgn2_6/ptodaca/1/paa/US099B_COMB.pep.*
- 27: /cgn2_6/ptodaca/1/paa/US100_COMB.pep.*
- 28: /cgn2_6/ptodaca/1/paa/US101_COMB.pep.*
- 29: /cgn2_6/ptodaca/1/paa/US102_COMB.pep.*
- 30: /cgn2_6/ptodaca/1/paa/US103_COMB.pep.*
- 31: /cgn2_6/ptodaca/1/paa/US104_COMB.pep.*
- 32: /cgn2_6/ptodaca/1/paa/US106_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	836	100.0	836 1	PCT-US02-05095A-1600 Sequence 1600, Ap

2	836	100.0	836	23	US-09-899-569a-4	Sequence 4, Appl
3	650	77.8	749	23	US-09-899-569a-2	Sequence 2, Appl
4	413	49.4	414	29	US-10-351-334-273	Sequence 273, App
5	413	49.4	443	1	PCT-US99-17130-265	Sequence 265, App
6	413	49.4	443	29	US-10-351-334-271	Sequence 271, App
7	382	45.7	649	20	US-09-629-469a-18191	Sequence 18191, A
8	341	40.8	343	1	PCT-US01-43523-162	Sequence 162, App
9	341	40.8	343	1	PCT-US02-24563-162	Sequence 162, App
10	341	40.8	343	26	US-10-028-072-162	Sequence 162, App
11	341	40.8	343	27	US-10-121-040-162	Sequence 162, App
12	341	40.8	343	27	US-10-121-041-162	Sequence 162, App
13	341	40.8	343	27	US-10-121-042-162	Sequence 162, App
14	341	40.8	343	27	US-10-121-043-162	Sequence 162, App
15	341	40.8	343	27	US-10-121-044-162	Sequence 162, App
16	341	40.8	343	27	US-10-121-045-162	Sequence 162, App
17	341	40.8	343	27	US-10-121-046-162	Sequence 162, App
18	341	40.8	343	27	US-10-121-047-162	Sequence 162, App
19	341	40.8	343	27	US-10-121-048-162	Sequence 162, App
20	341	40.8	343	27	US-10-121-049-162	Sequence 162, App
21	341	40.8	343	27	US-10-121-050-162	Sequence 162, App
22	341	40.8	343	27	US-10-121-051-162	Sequence 162, App
23	341	40.8	343	27	US-10-121-052-162	Sequence 162, App
24	341	40.8	343	27	US-10-121-053-162	Sequence 162, App
25	341	40.8	343	27	US-10-121-054-162	Sequence 162, App
26	341	40.8	343	27	US-10-121-055-162	Sequence 162, App
27	341	40.8	343	27	US-10-121-056-162	Sequence 162, App
28	341	40.8	343	27	US-10-121-057-162	Sequence 162, App
29	341	40.8	343	27	US-10-121-058-162	Sequence 162, App
30	341	40.8	343	27	US-10-121-059-162	Sequence 162, App
31	341	40.8	343	27	US-10-121-060-162	Sequence 162, App
32	341	40.8	343	27	US-10-121-061-162	Sequence 162, App
33	341	40.8	343	27	US-10-121-063-162	Sequence 162, App
34	341	40.8	343	27	US-10-123-108-162	Sequence 162, App
35	341	40.8	343	27	US-10-123-109-162	Sequence 162, App
36	341	40.8	343	27	US-10-123-154-162	Sequence 162, App
37	341	40.8	343	27	US-10-123-156-162	Sequence 162, App
38	341	40.8	343	27	US-10-123-157-162	Sequence 162, App
39	341	40.8	343	27	US-10-123-212-162	Sequence 162, App
40	341	40.8	343	27	US-10-123-213-162	Sequence 162, App
41	341	40.8	343	27	US-10-123-214-162	Sequence 162, App
42	341	40.8	343	27	US-10-123-215-162	Sequence 162, App
43	341	40.8	343	27	US-10-123-235-162	Sequence 162, App
44	341	40.8	343	27	US-10-123-236-162	Sequence 162, App
45	341	40.8	343	27	US-10-123-261-162	Sequence 162, App

ALIGNMENTS

RESULT 1
PCT-US02-05095A-1600
Sequence 1600, Application PC/TUS0205095A
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-075 (803)
CURRENT APPLICATION NUMBER: PCT/US02/05095A
PRIOR FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: 09/799,451
NUMBER OF SEQ ID NOS: 1896
SEQ ID NO 1600
LENGTH: 836
TYPE: PPT
ORGANISM: Homo sapiens
PCT-US02-05095A-1600

Query Match	100.0%	Score 836;	DB 1;	Length 836;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 836;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Oy	1	MAGNCGSIALLCVLLIGAARLPAGAAFPALPRESNTVTLTKGPTTLARPCVVI	60	

```
Db 1 MAGLNCGVSIALLGVLLIGAARLPKGAFAFIALPRESNITVLIKLGTPTLLAKPCYIV 60
Qy 61 SKRHITMLSIKSGERIVTFSCQSPNNHFVIEIQKNDICMSGPCPGEVQLOPSTSLPT 120
Db 61 SKRHITMLSIKSGERIVTFSCQSPNNHFVIEIQKNDICMSGPCPGEVQLOPSTSLPT 120
Qy 121 LNRFTFMDVKAHKSIGLELOFSIPRLROIIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
Db 121 LNRFTFMDVKAHKSIGLELOFSIPRLROIIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
Qy 181 GTVSRIKMOEGVKMALHLPWFHPRNVSGFSIANRSSIKRLCTIESVFEGESATLMSANY 240
Db 181 GTVSRIKMOEGVKMALHLPWFHPRNVSGFSIANRSSIKRLCTIESVFEGESATLMSANY 240
Qy 241 PEGFPEDELMTQOFVPAHLRASVSFLNLTNSCKEKEBEVEYIIGSTTNEVEFKLBDK 300
Db 241 PEGFPEDELMTQOFVPAHLRASVSFLNLTNSCKEKEBEVEYIIGSTTNEVEFKLBDK 300
Qy 301 QPGNAGNFNLSLOCCDDAOSPGILRLOFVLOHPONESNKIYVVDLSNERAMSLTIE 360
Db 301 QPGNAGNFNLSLOCCDDAOSPGILRLOFVLOHPONESNKIYVVDLSNERAMSLTIE 360
Qy 361 PRPVQSRKFVPGCFVCELSRTCSNLTLSGSKHISFLCDDLRLMANNVEKITSCTDH 420
Db 361 PRPVQSRKFVPGCFVCELSRTCSNLTLSGSKHISFLCDDLRLMANNVEKITSCTDH 420
Qy 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLLVPKORSLVLPKQKLOQTHEKPCNTSF 480
Db 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLLVPKORSLVLPKQKLOQTHEKPCNTSF 480
Qy 481 SYLVASAIPODLVFGSFCPGSGIKQIOVKONISVTLRTFAPSFOEASRQGLTVSFTPY 540
Db 481 SYLVASAIPODLVFGSFCPGSGIKQIOVKONISVTLRTFAPSFOEASRQGLTVSFTPY 540
Qy 541 FKBEGVFTVPTDTSKSVYLRTPNMDRGLPSLTSVSWNISVPRDOVAACLTFEKERSGVQC 600
Db 541 FKBEGVFTVPTDTSKSVYLRTPNMDRGLPSLTSVSWNISVPRDOVAACLTFEKERSGVQC 600
Qy 601 TGRAFMIIQEORTABEIEFSLDEEDVLPKPSFHHHSFWNISNCSPTSKQDLDFSVTLLT 660
Db 601 TGRAFMIIQEORTABEIEFSLDEEDVLPKPSFHHHSFWNISNCSPTSKQDLDFSVTLLT 660
Qy 661 PRVQKRSYSLOVPSDILHLPELHDFSWKLLVPKORSLVLPKQKLOQTHEKPCNTSF 480
Db 661 PRVQKRSYSLOVPSDILHLPELHDFSWKLLVPKORSLVLPKQKLOQTHEKPCNTSF 480
Qy 721 KFOKGRKNDSHVAVIEDTWYGHLLQDSSGSLQPEVDYTRPFOGTMGVCPSPPTIC 780
Db 721 KFOKGRKNDSHVAVIEDTWYGHLLQDSSGSLQPEVDYTRPFOGTMGVCPSPPTIC 780
Qy 781 SRAPTAKLATEBPPRSPSESESEPTTFSHPNNGDVSSKOTDIPLLNTQEBMEPAE 836
Db 781 SRAPTAKLATEBPPRSPSESESEPTTFSHPNNGDVSSKOTDIPLLNTQEBMEPAE 836

RESULT 2
US-09-899-569A-4
; Sequence 4, Application US/09899569A
; GENERAL INFORMATION:
; APPLICANT: Norbert Schweifer
; APPLICANT: Marwa Scherl-Mostageer
; APPLICANT: Wolfgang Sommergruber
; APPLICANT: Roger Abseher
; TITLE OF INVENTION: Tumorssoziierter Antigen (B345)
; FILE REFERENCE: 0653.2280001
; CURRENT APPLICATION NUMBER: US/09/899,569A
; CURRENT FILING DATE: 2001-07-06
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: DE 100 33 080.0
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 60/243,158
; PRIOR FILING DATE: 2000-10-25
```

```
; PRIOR APPLICATION NUMBER: US 60/297,747
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-899-569A-4

Query Match 100.0%; Score 836; DB 23; Length 836;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 836; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAGLNCGVSIALLGVLLIGAARLPKGAFAFIALPRESNITVLIKLGTPTLLAKPCYIV 60
Db 1 MAGLNCGVSIALLGVLLIGAARLPKGAFAFIALPRESNITVLIKLGTPTLLAKPCYIV 60
Qy 61 SKRHITMLSIKSGERIVTFSCQSPNNHFVIEIQKNDICMSGPCPGEVQLOPSTSLPT 120
Db 61 SKRHITMLSIKSGERIVTFSCQSPNNHFVIEIQKNDICMSGPCPGEVQLOPSTSLPT 120
Qy 121 LNRFTFMDVKAHKSIGLELOFSIPRLROIIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
Db 121 LNRFTFMDVKAHKSIGLELOFSIPRLROIIGPESCPDGVTHSISGRIDATVVRIGTFCSN 180
Qy 181 GTVSRIKMOEGVKMALHLPWFHPRNVSGFSIANRSSIKRLCTIESVFEGESATLMSANY 240
Db 181 GTVSRIKMOEGVKMALHLPWFHPRNVSGFSIANRSSIKRLCTIESVFEGESATLMSANY 240
Qy 241 PEGFPEDELMTQOFVPAHLRASVSFLNLTNSCKEKEBEVEYIIGSTTNEVEFKLBDK 300
Db 241 PEGFPEDELMTQOFVPAHLRASVSFLNLTNSCKEKEBEVEYIIGSTTNEVEFKLBDK 300
Qy 301 QPGNAGNFNLSLOCCDDAOSPGILRLOFVLOHPONESNKIYVVDLSNERAMSLTIE 360
Db 301 QPGNAGNFNLSLOCCDDAOSPGILRLOFVLOHPONESNKIYVVDLSNERAMSLTIE 360
Qy 361 PRPVQSRKFVPGCFVCELSRTCSNLTLSGSKHISFLCDDLRLMANNVEKITSCTDH 420
Db 361 PRPVQSRKFVPGCFVCELSRTCSNLTLSGSKHISFLCDDLRLMANNVEKITSCTDH 420
Qy 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLLVPKORSLVLPKQKLOQTHEKPCNTSF 480
Db 421 RYCQKRSYSLOVPSDILHLPELHDFSWKLLVPKORSLVLPKQKLOQTHEKPCNTSF 480
Qy 481 SYLVASAIPODLVFGSFCPGSGIKQIOVKONISVTLRTFAPSFOEASRQGLTVSFTPY 540
Db 481 SYLVASAIPODLVFGSFCPGSGIKQIOVKONISVTLRTFAPSFOEASRQGLTVSFTPY 540
Qy 541 FKBEGVFTVPTDTSKSVYLRTPNMDRGLPSLTSVSWNISVPRDOVAACLTFEKERSGVQC 600
Db 541 FKBEGVFTVPTDTSKSVYLRTPNMDRGLPSLTSVSWNISVPRDOVAACLTFEKERSGVQC 600
Qy 601 TGRAFMIIQEORTABEIEFSLDEEDVLPKPSFHHHSFWNISNCSPTSKQDLDFSVTLLT 660
Db 601 TGRAFMIIQEORTABEIEFSLDEEDVLPKPSFHHHSFWNISNCSPTSKQDLDFSVTLLT 660
Qy 661 PRVQKRSYSLOVPSDILHLPELHDFSWKLLVPKORSLVLPKQKLOQTHEKPCNTSF 480
Db 661 PRVQKRSYSLOVPSDILHLPELHDFSWKLLVPKORSLVLPKQKLOQTHEKPCNTSF 480
Qy 721 KFOKGRKNDSHVAVIEDTWYGHLLQDSSGSLQPEVDYTRPFOGTMGVCPSPPTIC 780
Db 721 KFOKGRKNDSHVAVIEDTWYGHLLQDSSGSLQPEVDYTRPFOGTMGVCPSPPTIC 780
Qy 781 SRAPTAKLATEBPPRSPSESESEPTTFSHPNNGDVSSKOTDIPLLNTQEBMEPAE 836
Db 781 SRAPTAKLATEBPPRSPSESESEPTTFSHPNNGDVSSKOTDIPLLNTQEBMEPAE 836

RESULT 3
US-09-899-569A-2
```

```
; Sequence 2, Application US/09899569A
; GENERAL INFORMATION:
; APPLICANT: Norbert Schweifer
; APPLICANT: Marwa Scherl-Mostager
; APPLICANT: Wolfgang Sommergruber
; APPLICANT: Roger Abbeher
; TITLE OF INVENTION: Tumorasoziiertes Antigen (B345)
; FILE REFERENCE: 0652.2280001
; CURRENT APPLICATION NUMBER: US/09/899,569A
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: DE 100 33 080.0
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: DE 101 19 294.0
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 60/243,158
; PRIOR FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: US 60/297,747
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 749
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-899-569A-2

Query Match      77.8%; Score 650; DB 23; Length 749;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 650; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 MLSTKGRIVFTFSCQSPENHFVIEIQKIDCMGSPCPFGVQLQPSSTLLPTLNRTFI 126
DB 1 MLSTKGRIVFTFSCQSPENHFVIEIQKIDCMGSPCPFGVQLQPSSTLLPTLNRTFI 60
QY 127 MDVAKHKSIGLELOPSIRLRQIGGSCPDGVTHTSIGRIDATVVRIGTFCSCNTVSR 186
DB 61 MDVAKHKSIGLELOPSIRLRQIGGSCPDGVTHTSIGRIDATVVRIGTFCSCNTVSR 120
QY 187 KMGEVKAHLPMFHPNVSQFSIANKRSIRKLCIISVFEGBSATLMSANYPEGPE 246
DB 121 KMGEVKAHLPMFHPNVSQFSIANKRSIRKLCIISVFEGBSATLMSANYPEGPE 180
QY 247 DELMTQFVPAHLRAVSFLNPNLSNCRKEEREVEYIIGSTNPEVFKLEDKOPGMA 306
DB 181 DELMTQFVPAHLRAVSFLNPNLSNCRKEEREVEYIIGSTNPEVFKLEDKOPGMA 240
QY 307 GNFNLSLQCCDDAOSPGLRLQFQVLVQHPONESNKIYVVDLSNERAMSLTIEPRVQ 366
DB 241 GNFNLSLQCCDDAOSPGLRLQFQVLVQHPONESNKIYVVDLSNERAMSLTIEPRVQ 300
QY 367 SRKFPVPGFVCLSESTCSSNLTLTSGSKHKSFLCDDLTRLMMNVEKTIISCTDHRVYCOR 426
DB 301 SRKFPVPGFVCLSESTCSSNLTLTSGSKHKSFLCDDLTRLMMNVEKTIISCTDHRVYCOR 360
QY 427 SYSLOVPSDILHLPEYLDHDFSWKLLVPRKRLSLVLPVPAKLOQHTHEPCNTSFSYLVA 486
DB 361 SYSLOVPSDILHLPEYLDHDFSWKLLVPRKRLSLVLPVPAKLOQHTHEPCNTSFSYLVA 420
QY 487 AIPSDOLYFGSFCPEGSIKQIOVKONISVTLRTFAPSFOEASRQGLTVSFIYKKEGV 546
DB 421 AIPSDOLYFGSFCPEGSIKQIOVKONISVTLRTFAPSFOEASRQGLTVSFIYKKEGV 480
QY 547 FTVPDTSKYVLRTPMNDRGLPSTLSVSNISVPRDVACTTFEKESGVVCCQGRAPM 606
DB 481 FTVPDTSKYVLRTPMNDRGLPSTLSVSNISVPRDVACTTFEKESGVVCCQGRAPM 540
QY 607 IIOEORTAEBSIFSLDEVLPRKPSFHHSFWNINSCSPSGKODLLFSVTLTPRTYDL 666
DB 541 IIOEORTAEBSIFSLDEVLPRKPSFHHSFWNINSCSPSGKODLLFSVTLTPRTYDL 600
QY 667 TVILIAAVGGVLLSALGLIICVYKXKKKKTKNGPAVGYNINTEMPT 716
DB 601 TVILIAAVGGVLLSALGLIICVYKXKKKKTKNGPAVGYNINTEMPT 650
```

```
RESULT 4
US-10-351-334-273
; Sequence 273, Application US/10351334
; GENERAL INFORMATION:
; APPLICANT: Komatsoulis et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P2
; CURRENT APPLICATION NUMBER: US/10/351,334
; PRIOR FILING DATE: 2003-01-27
; PRIOR APPLICATION NUMBER: 60/350,898
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/489,847
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/US99/17130
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/095,486
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/096,319
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: 60/095,454
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: 60/095,455
; PRIOR FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 273
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-351-334-273

Query Match      49.4%; Score 413; DB 29; Length 414;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGINCQVSTLLGLVLLGAARLRPGAEAFIALPRESENTIVLKLGTPTLLARPCYVI 60
DB 1 MAGINCQVSTLLGLVLLGAARLRPGAEAFIALPRESENTIVLKLGTPTLLARPCYVI 60
QY 61 SKRHITMLSTKGRIVFTFSCQSPENHFVIEIQKIDCMGSPCPFGVQLQPSSTLLPT 120
DB 61 SKRHITMLSTKGRIVFTFSCQSPENHFVIEIQKIDCMGSPCPFGVQLQPSSTLLPT 120
QY 121 LNRFTIMDVAKHKSIGLELOPSIRLRQIGGSCPDGVTHTSIGRIDATVVRIGTFCSCN 180
DB 121 LNRFTIMDVAKHKSIGLELOPSIRLRQIGGSCPDGVTHTSIGRIDATVVRIGTFCSCN 180
QY 181 GTVSRIRKQGEVKAHLPMFHPNVSQFSIANKRSIRKLCIISVFEGBSATLMSANY 240
DB 181 GTVSRIRKQGEVKAHLPMFHPNVSQFSIANKRSIRKLCIISVFEGBSATLMSANY 240
QY 241 PEGFPEDELMTQFVPAHLRAVSFLNPNLSNCRKEEREVEYIIGSTNPEVFKLEDK 300
DB 241 PEGFPEDELMTQFVPAHLRAVSFLNPNLSNCRKEEREVEYIIGSTNPEVFKLEDK 300
QY 301 QPGMAGNFNLSLQCCDDAOSPGLRLQFQVLVQHPONESNKIYVVDLSNERAMSLTIE 360
DB 301 QPGMAGNFNLSLQCCDDAOSPGLRLQFQVLVQHPONESNKIYVVDLSNERAMSLTIE 360
QY 361 PRPVKQSKFVPGFVCLSESTCSSNLTLTSGSKHKSFLCDDLTRLMMNVEK 413
DB 361 PRPVKQSKFVPGFVCLSESTCSSNLTLTSGSKHKSFLCDDLTRLMMNVEK 413

RESULT 5
PCT-US99-17130-265
; Sequence 265, Application PC/TUS9917130
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
```

TITLE OF INVENTION: 98 Human Secreted Proteins
FILE REFERENCE: P2031.PCT
CURRENT APPLICATION NUMBER: PCT/US99/17130
CURRENT FILING DATE: 1999-07-29
PRIOR APPLICATION NUMBER: 60/094,457
PRIOR FILING DATE: 1998-07-30
EARLIER APPLICATION NUMBER: 60/095,486
EARLIER FILING DATE: 1998-08-05
EARLIER APPLICATION NUMBER: 60/095,455
EARLIER FILING DATE: 1998-08-06
EARLIER APPLICATION NUMBER: 60/095,454
EARLIER FILING DATE: 1998-08-06
EARLIER APPLICATION NUMBER: 60/096,319
EARLIER FILING DATE: 1998-08-12
NUMBER OF SEQ ID NOS: 364
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 265
LENGTH: 443
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US99-17130-265

Query Match 49.4%; Score 413; DB 1; Length 443;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNGVSIALGLVLLGAARLPRGAFAFEIALPRESNITVLIKGTPTLLAKPCYIV 60
DB 30 MAGLNGVSIALGLVLLGAARLPRGAFAFEIALPRESNITVLIKGTPTLLAKPCYIV 89
QY 61 SKRHITMISIKGERIVFTFSCSPENHVIETIOKNIDMSGPCPGEVQLOPSTSLPT 120
DB 90 SKRHITMISIKGERIVFTFSCSPENHVIETIOKNIDMSGPCPGEVQLOPSTSLPT 149
QY 121 LNRTFIMDKAHKSIGLELOFSIPRLROIQPGSCPDGVTSHISGRIDATVIRIGTFCSN 180
DB 150 LNRTFIMDKAHKSIGLELOFSIPRLROIQPGSCPDGVTSHISGRIDATVIRIGTFCSN 209
QY 181 GTVSRIKQEGVKALHPMFHPRNVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 240
DB 210 GTVSRIKQEGVKALHPMFHPRNVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 269
QY 241 PGGFPEDELTMTQFVVPALRASVSFLNPNLSNCKEKREVEYYIGSTTNEVEFLBCK 300
DB 270 PGGFPEDELTMTQFVVPALRASVSFLNPNLSNCKEKREVEYYIGSTTNEVEFLBCK 329
QY 301 QPGNAGNFNLISLQCDQDQSPGILRLQFVLYVHPONESNKIYVVDLSNERAMSLTIE 360
DB 330 QPGNAGNFNLISLQCDQDQSPGILRLQFVLYVHPONESNKIYVVDLSNERAMSLTIE 389
QY 361 PRPVKSRKFPVPGCFVCLSESRTCSSNLTITSGSKHISFLCDLTRLMMNVK 413
DB 390 PRPVKSRKFPVPGCFVCLSESRTCSSNLTITSGSKHISFLCDLTRLMMNVK 442

RESULT 6
US-10-351-334-271
Sequence 271, Application US/10351334

GENERAL INFORMATION:

APPLICANT: Komatsushita et al

TITLE OF INVENTION: 98 Human Secreted Proteins

FILE REFERENCE: P2031P2

CURRENT APPLICATION NUMBER: US/10/351,334

CURRENT FILING DATE: 2003-01-27

PRIOR APPLICATION NUMBER: 60/350,898

PRIOR FILING DATE: 2002-01-25

PRIOR APPLICATION NUMBER: 09/489,847

PRIOR FILING DATE: 2000-01-24

PRIOR APPLICATION NUMBER: PCT/US99/17130

PRIOR FILING DATE: 1999-07-29

PRIOR APPLICATION NUMBER: 60/094,657

PRIOR FILING DATE: 1998-07-30

PRIOR APPLICATION NUMBER: 60/095,486

PRIOR FILING DATE: 1998-08-05
PRIOR APPLICATION NUMBER: 60/096,319
PRIOR FILING DATE: 1998-08-12
PRIOR APPLICATION NUMBER: 60/095,454
PRIOR FILING DATE: 1998-08-06
PRIOR APPLICATION NUMBER: 60/095,455
PRIOR FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 376
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 271
LENGTH: 443
TYPE: PRT
ORGANISM: Homo sapiens
US-10-351-334-271

Query Match 49.4%; Score 413; DB 29; Length 443;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNGVSIALGLVLLGAARLPRGAFAFEIALPRESNITVLIKGTPTLLAKPCYIV 60
DB 30 MAGLNGVSIALGLVLLGAARLPRGAFAFEIALPRESNITVLIKGTPTLLAKPCYIV 89
QY 61 SKRHITMISIKGERIVFTFSCSPENHVIETIOKNIDMSGPCPGEVQLOPSTSLPT 120
DB 90 SKRHITMISIKGERIVFTFSCSPENHVIETIOKNIDMSGPCPGEVQLOPSTSLPT 149
QY 121 LNRTFIMDKAHKSIGLELOFSIPRLROIQPGSCPDGVTSHISGRIDATVIRIGTFCSN 180
DB 150 LNRTFIMDKAHKSIGLELOFSIPRLROIQPGSCPDGVTSHISGRIDATVIRIGTFCSN 209
QY 181 GTVSRIKQEGVKALHPMFHPRNVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 240
DB 210 GTVSRIKQEGVKALHPMFHPRNVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 269
QY 241 PGGFPEDELTMTQFVVPALRASVSFLNPNLSNCKEKREVEYYIGSTTNEVEFLBCK 300
DB 270 PGGFPEDELTMTQFVVPALRASVSFLNPNLSNCKEKREVEYYIGSTTNEVEFLBCK 329
QY 301 QPGNAGNFNLISLQCDQDQSPGILRLQFVLYVHPONESNKIYVVDLSNERAMSLTIE 360
DB 330 QPGNAGNFNLISLQCDQDQSPGILRLQFVLYVHPONESNKIYVVDLSNERAMSLTIE 389
QY 361 PRPVKSRKFPVPGCFVCLSESRTCSSNLTITSGSKHISFLCDLTRLMMNVK 413
DB 390 PRPVKSRKFPVPGCFVCLSESRTCSSNLTITSGSKHISFLCDLTRLMMNVK 442

RESULT 7
US-09-629-469A-18191
Sequence 18191, Application US/09629469A

GENERAL INFORMATION:

APPLICANT: OTA, TOSHIO

APPLICANT: ISOGAI, TAKAO

APPLICANT: NISHIKAWA, TETSUO

APPLICANT: HAYASHI, KOJI

APPLICANT: SAITO, KAORU

APPLICANT: YAMAMOTO, JUNICHI

APPLICANT: ISHII, SHIZUKO

APPLICANT: SUGIYAMA, TOMOYASU

APPLICANT: WAKAMATSU, AI

APPLICANT: NAGAI, KEIICHI

APPLICANT: OTSUKI, TETSUJI

TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE

FILE REFERENCE: 084335/0123

CURRENT APPLICATION NUMBER: US/09/629,469A

CURRENT FILING DATE: 2000-07-28

PRIOR APPLICATION NUMBER: JP 1999-248036

PRIOR FILING DATE: 1999-07-29

PRIOR APPLICATION NUMBER: JP 1999-300253

PRIOR FILING DATE: 1999-08-27

PRIOR APPLICATION NUMBER: JP 2000-118776

PRIOR FILING DATE: 2000-01-11

PRIOR APPLICATION NUMBER: JP 2000-183767
PRIOR FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: JP 2000-241899
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/159,590
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 60/183,322
PRIOR FILING DATE: 2000-02-17
NUMBER OF SEQ ID NOS: 19025
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 18191
LENGTH: 649
TYPE: PRT
ORGANISM: Homo sapiens
US-09-629-469A-18191

Query Match 45.7%; Score 382; DB 20; Length 649;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 582; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

253 QFVPAHLRASVFLNPNLNCERKEERYIIPGTTPEVFKLEDKOPGMAGNPNLS 312
66 QFVPAHLRASVFLNPNLNCERKEERYIIPGTTPEVFKLEDKOPGMAGNPNLS 125
313 LOGCDQDASPGILRLQFVLVQHPONESNKIYVVDLSNERAMSLLTEBRPVKQSRKFPV 372
126 LOGCDQDASPGILRLQFVLVQHPONESNKIYVVDLSNERAMSLLTEBRPVKQSRKFPV 185
373 GCFVLESTGSSNLTLSGSGHKSIFLCLDITLIMANVEKTIISTDHRVYCKRKSYSLOY 432
186 GCFVLESTGSSNLTLSGSGHKSIFLCLDITLIMANVEKTIISTDHRVYCKRKSYSLOY 245
433 PSDIHLPEVLEDFGSKLIVPKDRLSLVLPAPQKLOQHTHEKPCNTSFSYVASAIPSD 492
246 PSDIHLPEVLEDFGSKLIVPKDRLSLVLPAPQKLOQHTHEKPCNTSFSYVASAIPSD 305
493 LYFGSFCPGSIIKOIVKONISVTLRTFAPSFOQASROGLTVSFIYFKERGVFTVTPD 552
306 LYFGSFCPGSIIKOIVKONISVTLRTFAPSFOQASROGLTVSFIYFKERGVFTVTPD 365
553 TKSATYLLTPMNDRLGSLITSVSNISVPRDOVACLTFFKERSGVVCCOTGRAFMIIQEQ 612
366 TKSATYLLTPMNDRLGSLITSVSNISVPRDOVACLTFFKERSGVVCCOTGRAFMIIQEQ 425
613 TRAEIFSLDEVDLKPSPFHHSFMWNISNCSPTSGKQDLIFSVTLIPRTVDLVIIILA 672
426 TRAEIFSLDEVDLKPSPFHHSFMWNISNCSPTSGKQDLIFSVTLIPRTVDLVIIILA 485
673 AVGGVLLLSALGLIICVKKKKKKTKNGPAVGIYNGNINTEMPROPKKFQGRKDNDSH 732
486 AVGGVLLLSALGLIICVKKKKKKTKNGPAVGIYNGNINTEMPROPKKFQGRKDNDSH 545
733 YVAVIEDTMVYGHLLQDSSGSFLOPEVDVTPRPQGTGVCPPSPPTIGSRAPTAKLATEE 792
546 YVAVIEDTMVYGHLLQDSSGSFLOPEVDVTPRPQGTGVCPPSPPTIGSRAPTAKLATEE 605
793 PPPRSPSESEPRYTFSPHNGDVSSKOTDIPLNTQEMEPAE 836
606 PPPRSPSESEPRYTFSPHNGDVSSKOTDIPLNTQEMEPAE 649

RESULT 8

PCT-US01-43523-162
Sequence 162, Application PC/TUS0143523
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerltzen, Mary E.
APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zhen
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C31
CURRENT APPLICATION NUMBER: PCT/US01/43523
CURRENT FILING DATE: 2002-10-02
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 162
LENGTH: 343
TYPE: PRT
ORGANISM: Homo Sapien
PCT-US01-43523-162

Query Match 40.8%; Score 341; DB 1; Length 343;
Best Local Similarity 100.0%; Pred. No. 36-304;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MAGLNGVSIALGLVLLGAARLPGRAPAFIALPRESNITVLKLGPTLLAKPCYIVI 60
1 MAGLNGVSIALGLVLLGAARLPGRAPAFIALPRESNITVLKLGPTLLAKPCYIVI 60
61 SKRHTMISTSGRIVTTFSCQSPENHFVIEIQKNIDCMGPGCFGVOLQPSISLPT 120
61 SKRHTMISTSGRIVTTFSCQSPENHFVIEIQKNIDCMGPGCFGVOLQPSISLPT 120
121 LNRFTIMVKAHKSIGLEQSPILRLQIGPESCPDVTHSISGRIDATVIRIGTFCSN 180
121 LNRFTIMVKAHKSIGLEQSPILRLQIGPESCPDVTHSISGRIDATVIRIGTFCSN 180
181 GTVSRIKQEGVKALHLPMFPRNVSGFSIANSSIRLCTISVFEGBSATLMSANY 240
181 GTVSRIKQEGVKALHLPMFPRNVSGFSIANSSIRLCTISVFEGBSATLMSANY 240
241 PEGFPEDILMTWQFVPAHLRASVFLNPNLNCERKEERYIIPGTTPEVFKLEDK 300
241 PEGFPEDILMTWQFVPAHLRASVFLNPNLNCERKEERYIIPGTTPEVFKLEDK 300
301 QPGMAGNPNLSLOGCDQDASPGILRLQFVLVQHPONES 341
301 QPGMAGNPNLSLOGCDQDASPGILRLQFVLVQHPONES 341

RESULT 9

PCT-US02-24563-162

```
Sequence 162, Application PC/TUS0224563
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330RLC31
CURRENT APPLICATION NUMBER: PCT/US02/24563
PRIOR FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION data removed - See File Wrapper or PAM.
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 162
LENGTH: 343
TYPE: PRT
ORGANISM: Homo Sapien
PCT-US02-24563-162

Query Match      40.8%; Score 341; DB 1; Length 343;
Best Local Similarity 100.0%; Pred No. 3e-304;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Db      241 PEGPEDELMTWQFVPAHLRASVSFLNPNLSNCKEERVEYYIPGSTTNPVEFKLEDK 300
Qy      301 QPGNAGNPNLSLQGCDDAOSPGILRLQFVLYVHPONES 341
Db      301 QPGNAGNPNLSLQGCDDAOSPGILRLQFVLYVHPONES 341

RESULT 10
US-10-028-072-162
Sequence 162, Application US/10028072
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION:
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/028,072
PRIOR FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059836
PRIOR FILING DATE: 1997-09-24
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062285
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062814
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/062816
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063045
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063082
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/063127
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063327
PRIOR FILING DATE: 1997-10-27
```


PRIOR APPLICATION NUMBER: 60/063329
 PRIOR FILING DATE: 1997-10-27
 PRIOR APPLICATION NUMBER: 60/063350
 PRIOR FILING DATE: 1997-10-28
 PRIOR APPLICATION NUMBER: 60/063561
 PRIOR FILING DATE: 1997-10-28
 PRIOR APPLICATION NUMBER: 60/063704
 PRIOR FILING DATE: 1997-10-29
 PRIOR APPLICATION NUMBER: 60/063733
 PRIOR FILING DATE: 1997-10-29
 PRIOR APPLICATION NUMBER: 60/063735
 PRIOR FILING DATE: 1997-10-29
 PRIOR APPLICATION NUMBER: 60/063738
 PRIOR FILING DATE: 1997-10-29
 PRIOR APPLICATION NUMBER: 60/063755
 PRIOR FILING DATE: 1997-10-17
 PRIOR APPLICATION NUMBER: 60/064248
 PRIOR FILING DATE: 1997-11-03
 PRIOR APPLICATION NUMBER: 60/064809
 PRIOR FILING DATE: 1997-11-07
 PRIOR APPLICATION NUMBER: 60/065186
 PRIOR FILING DATE: 1997-11-12
 PRIOR APPLICATION NUMBER: 60/065846
 PRIOR FILING DATE: 1997-11-17
 PRIOR APPLICATION NUMBER: 60/066364
 PRIOR FILING DATE: 1997-11-21
 PRIOR APPLICATION NUMBER: 60/066453
 PRIOR FILING DATE: 1997-11-24
 PRIOR APPLICATION NUMBER: 60/066511
 PRIOR FILING DATE: 1997-11-24
 PRIOR APPLICATION NUMBER: 60/066770
 PRIOR FILING DATE: 1997-11-24
 PRIOR APPLICATION NUMBER: 60/069212
 PRIOR FILING DATE: 1997-12-11
 PRIOR APPLICATION NUMBER: 60/069278
 PRIOR FILING DATE: 1997-12-11
 PRIOR APPLICATION NUMBER: 60/069334
 PRIOR FILING DATE: 1997-12-11
 PRIOR APPLICATION NUMBER: 60/069694
 PRIOR FILING DATE: 1997-12-16
 PRIOR APPLICATION NUMBER: 60/072320
 PRIOR FILING DATE: 1998-01-23
 PRIOR APPLICATION NUMBER: 60/073612
 PRIOR FILING DATE: 1998-02-04
 PRIOR APPLICATION NUMBER: 60/074086
 PRIOR FILING DATE: 1998-02-09
 PRIOR APPLICATION NUMBER: 60/074092
 PRIOR FILING DATE: 1998-02-09
 PRIOR APPLICATION NUMBER: 60/077791
 PRIOR FILING DATE: 1998-03-12
 PRIOR APPLICATION NUMBER: 60/078910
 PRIOR FILING DATE: 1998-03-20
 PRIOR APPLICATION NUMBER: 60/079294
 PRIOR FILING DATE: 1998-03-25
 PRIOR APPLICATION NUMBER: 60/079663
 PRIOR FILING DATE: 1998-02-27
 PRIOR APPLICATION NUMBER: 60/079728
 PRIOR FILING DATE: 1998-03-27
 PRIOR APPLICATION NUMBER: 60/080165
 PRIOR FILING DATE: 1998-03-31
 PRIOR APPLICATION NUMBER: 60/081203
 PRIOR FILING DATE: 1998-04-09
 PRIOR APPLICATION NUMBER: 60/081229
 PRIOR FILING DATE: 1998-04-09
 PRIOR APPLICATION NUMBER: 60/081695
 PRIOR FILING DATE: 1998-04-14
 PRIOR APPLICATION NUMBER: 60/081817
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/081818
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/082999
 PRIOR FILING DATE: 1998-04-24
 PRIOR APPLICATION NUMBER: 60/083322

PRIOR FILING DATE: 1998-04-28
 PRIOR APPLICATION NUMBER: 60/083545
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/084600
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084627
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084637
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/085149
 PRIOR FILING DATE: 1998-05-12
 PRIOR APPLICATION NUMBER: 60/085323
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085338
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085339
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085579
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085697
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085704
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/086414
 PRIOR FILING DATE: 1998-05-22
 PRIOR APPLICATION NUMBER: 60/086430
 PRIOR FILING DATE: 1998-05-22
 PRIOR APPLICATION NUMBER: 60/087106
 PRIOR FILING DATE: 1998-05-28
 PRIOR APPLICATION NUMBER: 60/088026
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088730
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088741
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088810
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088858
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/089532
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089599
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089907
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089947
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/090349
 PRIOR FILING DATE: 1998-06-23
 PRIOR APPLICATION NUMBER: 60/090429
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090445
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090538
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090863
 PRIOR FILING DATE: 1998-06-26
 PRIOR APPLICATION NUMBER: 60/091360
 PRIOR FILING DATE: 1998-07-01
 PRIOR APPLICATION NUMBER: 60/091519
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091982
 PRIOR FILING DATE: 1998-07-07
 PRIOR APPLICATION NUMBER: 60/093339

Query Match 40.8%; Score 341; DB 26; Length 343;
 Best Local Similarity 100.0%; Pred. No. 3e-304;
 Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLGAARLPGGAFAFIALPRESNITVLLIKGTPTLLAKPCYIVI 60
 DB 1 MAGLNCGVSIALLGVLLGAARLPGGAFAFIALPRESNITVLLIKGTPTLLAKPCYIVI 60

QY 61 SKRHITMISIKSGERIVTFPSCSPENHVFIEIQKIDCMGSPCPGEVQLOPSTSLPT 120
DB 61 SKRHITMISIKSGERIVTFPSCSPENHVFIEIQKIDCMGSPCPGEVQLOPSTSLPT 120
QY 121 LNRFTIWDVKAHKSIGLEIQFSIPRLROIQPGESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIWDVKAHKSIGLEIQFSIPRLROIQPGESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKMOEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 240
DB 181 GTVSRIKMOEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 240
QY 241 PEGFPEDELMTQFVVPVPAHLRASVSFLNPNLSNCRKEERVEYYIPGSTTNEVFKLEBK 300
DB 241 PEGFPEDELMTQFVVPVPAHLRASVSFLNPNLSNCRKEERVEYYIPGSTTNEVFKLEBK 300
QY 301 QPGNMGNFNLSLQCGDDAOSPGILRLQFOVLVOHPONES 341
DB 301 QPGNMGNFNLSLQCGDDAOSPGILRLQFOVLVOHPONES 341

RESULT 11

US-10-121-040-162
Sequence 162, Application US/10121040

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C7
CURRENT APPLICATION NUMBER: US/10/121,040
CURRENT FILING DATE: 2002-04-11
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 162
LENGTH: 343
TYPE: PRT
ORGANISM: Homo Sapien
US-10-121-040-162

Query Match 40.8%; Score 341; DB 27; Length 343;

Best Local Similarity 100.0%; Pred. No. 3e-304;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLGAARLPFGAFAFEIALPRESNITVLIKGTPTLLAKPCYIYI 60
DB 1 MAGLNCGVSIALLGVLLGAARLPFGAFAFEIALPRESNITVLIKGTPTLLAKPCYIYI 60
QY 61 SKRHITMISIKSGERIVTFPSCSPENHVFIEIQKIDCMGSPCPGEVQLOPSTSLPT 120
DB 61 SKRHITMISIKSGERIVTFPSCSPENHVFIEIQKIDCMGSPCPGEVQLOPSTSLPT 120
QY 121 LNRFTIWDVKAHKSIGLEIQFSIPRLROIQPGESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIWDVKAHKSIGLEIQFSIPRLROIQPGESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKMOEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 240
DB 181 GTVSRIKMOEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 240

DB 181 GTVSRIKMOEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 240
QY 241 PEGFPEDELMTQFVVPVPAHLRASVSFLNPNLSNCRKEERVEYYIPGSTTNEVFKLEBK 300
DB 241 PEGFPEDELMTQFVVPVPAHLRASVSFLNPNLSNCRKEERVEYYIPGSTTNEVFKLEBK 300
QY 301 QPGNMGNFNLSLQCGDDAOSPGILRLQFOVLVOHPONES 341
DB 301 QPGNMGNFNLSLQCGDDAOSPGILRLQFOVLVOHPONES 341

RESULT 12

US-10-121-041-162
Sequence 162, Application US/10121041

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C9
CURRENT APPLICATION NUMBER: US/10/121,041
CURRENT FILING DATE: 2002-04-11
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 162
LENGTH: 343
TYPE: PRT
ORGANISM: Homo Sapien
US-10-121-041-162

Query Match 40.8%; Score 341; DB 27; Length 343;

Best Local Similarity 100.0%; Pred. No. 3e-304;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNCGVSIALLGVLLGAARLPFGAFAFEIALPRESNITVLIKGTPTLLAKPCYIYI 60
DB 1 MAGLNCGVSIALLGVLLGAARLPFGAFAFEIALPRESNITVLIKGTPTLLAKPCYIYI 60
QY 61 SKRHITMISIKSGERIVTFPSCSPENHVFIEIQKIDCMGSPCPGEVQLOPSTSLPT 120
DB 61 SKRHITMISIKSGERIVTFPSCSPENHVFIEIQKIDCMGSPCPGEVQLOPSTSLPT 120
QY 121 LNRFTIWDVKAHKSIGLEIQFSIPRLROIQPGESCPDGVTHSISGRIDATVVRIGTFCSN 180
DB 121 LNRFTIWDVKAHKSIGLEIQFSIPRLROIQPGESCPDGVTHSISGRIDATVVRIGTFCSN 180
QY 181 GTVSRIKMOEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 240
DB 181 GTVSRIKMOEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESEVFEBSGATLMSANY 240
QY 241 PEGFPEDELMTQFVVPVPAHLRASVSFLNPNLSNCRKEERVEYYIPGSTTNEVFKLEBK 300
DB 241 PEGFPEDELMTQFVVPVPAHLRASVSFLNPNLSNCRKEERVEYYIPGSTTNEVFKLEBK 300
QY 301 QPGNMGNFNLSLQCGDDAOSPGILRLQFOVLVOHPONES 341
DB 301 QPGNMGNFNLSLQCGDDAOSPGILRLQFOVLVOHPONES 341

APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C5
CURRENT APPLICATION NUMBER: US/10/121,044
PRIORITY DATE: 2002-04-11
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 162
LENGTH: 343
TYPE: PRT
ORGANISM: Homo Sapien
US-10-121-044-162

Query Match 40.8%; Score 341; DB 27; Length 343;
Best Local Similarity 100.0%; Pred. No. 3e-304;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGLNGVSIALGLVLLGAARLPAGAAFEIALPRESNITVLIKLGPTLLAKPCYIVI 60
DB 1 MAGLNGVSIALGLVLLGAARLPAGAAFEIALPRESNITVLIKLGPTLLAKPCYIVI 60
QY 61 SKRHITMLSIKSGERTVFTFSCQSPENHFVIEIQKNIDCMGSPCPGEVQLQPTSLPT 120
DB 61 SKRHITMLSIKSGERTVFTFSCQSPENHFVIEIQKNIDCMGSPCPGEVQLQPTSLPT 120
QY 121 LNRFTIWDYKAHKSIGLELOFSIPRLQIGPGECPDGVTHSISGRIDATVIRIGTFCSN 180
DB 121 LNRFTIWDYKAHKSIGLELOFSIPRLQIGPGECPDGVTHSISGRIDATVIRIGTFCSN 180
QY 181 GTVSRIKMQEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESVFESEGSATLMSANY 240
DB 181 GTVSRIKMQEGVMALHLPWFHPRNVSGFSIANRSSIKRLCTIESVFESEGSATLMSANY 240
QY 241 PEGFPDELTMTQFVVPALHRASVSFLNLSNCKEKREVEYITPGSTTNPEVFTLEDK 300
DB 241 PEGFPDELTMTQFVVPALHRASVSFLNLSNCKEKREVEYITPGSTTNPEVFTLEDK 300
QY 301 QPGNMAGNFNLISLQGCDDAOSPGILRLQFOVLVQHONES 341
DB 301 QPGNMAGNFNLISLQGCDDAOSPGILRLQFOVLVQHONES 341

Search completed: February 20, 2004, 19:59:03
Job time : 363 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: February 20, 2004, 19:49:29 ; Search time 44 Seconds
(without alignments)
1495.299 Million cell updates/sec

Title: US-09-899-569A-4
Perfect score: 836
Sequence: 1 MAGINCGVSIALLGVLLGA.....SSKDTDIPLNTQEBMEBAE 836

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 378584 seqs, 7869973 residues

Word size : 0

Total number of hits satisfying chosen parameters: 378584

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database : Pending_Patents_AA_New:*

1: /cgn2_6/ptocdata/1/paa/PCT_NEW_COMB.pep:*\n2: /cgn2_6/ptocdata/1/paa/US06_NEW_COMB.pep:*\n3: /cgn2_6/ptocdata/1/paa/US07_NEW_COMB.pep:*\n4: /cgn2_6/ptocdata/1/paa/US08_NEW_COMB.pep:*\n5: /cgn2_6/ptocdata/1/paa/US09_NEW_COMB.pep:*\n6: /cgn2_6/ptocdata/1/paa/US10_NEW_COMB.pep:*\n7: /cgn2_6/ptocdata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	1.1	293	6	US-10-481-265-105
2	8	1.0	134	1	PCT-US03-38896-178
3	8	1.0	147	6	US-10-669-657-162
4	8	1.0	187	1	PCT-US03-38829-199
5	8	1.0	193	1	PCT-US03-38896-177
6	8	1.0	221	5	US-09-978-360A-436
7	8	1.0	221	6	US-10-679-063-6997
8	8	1.0	221	6	US-10-679-063-6864
9	8	1.0	221	6	US-10-679-063-11202
10	8	1.0	222	6	US-10-425-114A-50595
11	8	1.0	276	6	US-10-425-114A-11744
12	8	1.0	348	6	US-10-739-930-8397
13	8	1.0	377	1	PCT-US03-19153-296
14	8	1.0	377	6	US-10-463-720-296
15	8	1.0	485	6	US-10-425-114A-65537
16	8	1.0	487	6	US-10-425-114A-64936
17	8	1.0	499	6	US-10-425-114A-40938
18	8	1.0	499	6	US-10-425-114A-42517
19	8	1.0	631	5	US-09-614-150A-26727
20	8	1.0	634	6	US-10-425-114A-59648
21	7	0.8	19	5	US-09-341-590A-122
22	7	0.8	30	6	US-10-296-734-714
23	7	0.8	30	6	US-10-296-734-716
24	7	0.8	61	6	US-10-679-063-24192
25	7	0.8	75	6	US-10-705-531-4
26	7	0.8	77	5	US-09-796-692A-1805

27	7	0.8	78	6	US-10-415-182A-10176	Sequence 10176, A
28	7	0.8	79	6	US-10-415-182A-2392	Sequence 2392, Ap
29	7	0.8	82	6	US-10-425-114A-45433	Sequence 45433, A
30	7	0.8	82	6	US-10-462-691-29	Sequence 29, Appl
31	7	0.8	83	5	US-09-614-150A-32895	Sequence 32895, A
32	7	0.8	90	6	US-10-425-114A-58082	Sequence 58082, A
33	7	0.8	97	6	US-10-460-594-94	Sequence 94, Appl
34	7	0.8	113	6	US-10-679-063-18469	Sequence 18469, A
35	7	0.8	129	6	US-10-425-114A-67280	Sequence 67280, A
36	7	0.8	140	6	US-10-425-114A-58799	Sequence 58799, A
37	7	0.8	146	6	US-10-679-063-5291	Sequence 5291, Ap
38	7	0.8	157	6	US-10-425-114A-63652	Sequence 63652, A
39	7	0.8	165	6	US-10-679-063-4799	Sequence 4799, Ap
40	7	0.8	169	6	US-10-679-063-5043	Sequence 5043, Ap
41	7	0.8	176	6	US-10-679-063-4736	Sequence 4736, Ap
42	7	0.8	177	6	US-10-425-114A-46400	Sequence 46400, A
43	7	0.8	177	6	US-10-425-114A-73004	Sequence 73004, A
44	7	0.8	179	6	US-10-425-114A-49322	Sequence 49322, A
45	7	0.8	185	6	US-10-425-114A-53476	Sequence 53476, A

ALIGNMENTS

RESULT 1
US-10-481-265-105
Sequence 105: Application US/10481265
GENERAL INFORMATION:
APPLICANT: James, Brian William
APPLICANT: Bacon, Joanna
TITLE OF INVENTION: Mycobacterial Antigens Expressed under Low Oxygen Tension
FILE REFERENCE: 1581.1020000
CURRENT APPLICATION NUMBER: US/10/481,265
PRIOR FILING DATE: 2003-12-19
PRIOR APPLICATION NUMBER: GB 0115365.9
PRIOR FILING DATE: 2001-06-22
PRIOR APPLICATION NUMBER: GB 0121780.1
PRIOR FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: PCT/GB02/02845
PRIOR FILING DATE: 2002-06-21
NUMBER OF SEQ ID NOS: 138
SOFTWARE: PatentIn version 3.1
SEQ ID NO 105
LENGTH: 293
TYPE: PRT
ORGANISM: Mycobacterium tuberculosis
US-10-481-265-105
Query Match 1.1%; Score 9; DB 6; Length 293;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 674 VGGGVLLLS 682
67 VGGGVLLLS 75
RESULT 2
PCT-US03-38896-178
Sequence 178: Application PC/TUS0338896
GENERAL INFORMATION:
APPLICANT: diadexue, Inc.
APPLICANT: Macina, Roberto
APPLICANT: Turner, Leah
APPLICANT: Sun, Yongming
APPLICANT: Liu, Shu-hui
TITLE OF INVENTION: Compositions, Splice Variants and Methods Relating to Lung Specific
FILE REFERENCE: DEX-0454
CURRENT APPLICATION NUMBER: PCT/US03/38896
CURRENT FILING DATE: 2003-12-08
PRIOR APPLICATION NUMBER: US 60/431,516

; PRIOR FILING DATE: 2002-12-06
 ; PRIOR APPLICATION NUMBER: US 60/431,510
 ; PRIOR FILING DATE: 2002-12-06
 ; PRIOR APPLICATION NUMBER: US 60/431,307
 ; PRIOR FILING DATE: 2002-12-06
 ; NUMBER OF SEQ ID NOS: 279
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 178
 ; LENGTH: 134
 ; TYPE: PRT
 ; ORGANISM: Homo sapien
 PCT-US03-38896-178

Query Match 1.0%; Score 8; DB 1; Length 134;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 693 KKKKKTNK 700
 DB 70 KKKKKTNK 77

RESULT 3
 US-10-669-657-162
 ; Sequence 162, Application US/10669657
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen, et. al
 ; TITLE OF INVENTION: 124 Human Secreted Proteins
 ; FILE REFERENCE: PS956
 ; CURRENT APPLICATION NUMBER: US/10/669,657
 ; CURRENT FILING DATE: 2003-09-25
 ; PRIOR APPLICATION NUMBER: PCT/US02/09135
 ; PRIOR FILING DATE: 2002-03-26
 ; PRIOR APPLICATION NUMBER: 10/105,299
 ; PRIOR FILING DATE: 2002-03-26
 ; PRIOR APPLICATION NUMBER: 60/278,650
 ; PRIOR FILING DATE: 2001-03-27
 ; PRIOR APPLICATION NUMBER: 09/950,082
 ; PRIOR FILING DATE: 2001-09-12
 ; PRIOR APPLICATION NUMBER: 09/950,083
 ; PRIOR FILING DATE: 2001-09-12
 ; PRIOR APPLICATION NUMBER: 09/833,245
 ; PRIOR FILING DATE: 2001-04-12
 ; PRIOR APPLICATION NUMBER: PCT/US01/11988
 ; PRIOR FILING DATE: 2001-04-12
 ; PRIOR APPLICATION NUMBER: 60/331,287
 ; PRIOR FILING DATE: 2001-11-13
 ; PRIOR APPLICATION NUMBER: 60/277,340
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/306,171
 ; PRIOR FILING DATE: 2001-07-19
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 493
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 162
 ; LENGTH: 147
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-669-657-162

Query Match 1.0%; Score 8; DB 6; Length 147;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LIGVLLG 19
 DB 30 LIGVLLG 37

RESULT 4
 PCT-US03-38829-199
 ; Sequence 199, Application PC/TUS0338829
 ; GENERAL INFORMATION:

; APPLICANT: diadexus, Inc.
 ; APPLICANT: Macina, Roberto
 ; APPLICANT: Turner, Leah
 ; APPLICANT: Sun, Yongming
 ; TITLE OF INVENTION: Compositions, Splice Variants and Methods Relating to Breast Spec
 ; FILE REFERENCE: DEX-0453
 ; CURRENT APPLICATION NUMBER: PCT/US03/38829
 ; CURRENT FILING DATE: 2003-12-05
 ; PRIOR APPLICATION NUMBER: US 60/431,145
 ; PRIOR FILING DATE: 2002-12-05
 ; NUMBER OF SEQ ID NOS: 253
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 199
 ; LENGTH: 187
 ; TYPE: PRT
 ; ORGANISM: Homo sapien
 PCT-US03-38829-199

Query Match 1.0%; Score 8; DB 1; Length 187;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LIGVLLG 19
 DB 70 LIGVLLG 77

RESULT 5
 PCT-US03-38896-177
 ; Sequence 177, Application PC/TUS0338896
 ; GENERAL INFORMATION:
 ; APPLICANT: diadexus, Inc.
 ; APPLICANT: Macina, Roberto
 ; APPLICANT: Turner, Leah
 ; APPLICANT: Sun, Yongming
 ; APPLICANT: Liu, Shu-Hui
 ; TITLE OF INVENTION: Compositions, Splice Variants and Methods Relating to Lung Specifi
 ; FILE REFERENCE: DEX-0454
 ; CURRENT APPLICATION NUMBER: PCT/US03/38896
 ; CURRENT FILING DATE: 2003-12-08
 ; PRIOR APPLICATION NUMBER: US 60/431,516
 ; PRIOR FILING DATE: 2002-12-06
 ; PRIOR APPLICATION NUMBER: US 60/431,510
 ; PRIOR FILING DATE: 2002-12-06
 ; PRIOR APPLICATION NUMBER: 60/431,307
 ; PRIOR FILING DATE: 2002-12-06
 ; NUMBER OF SEQ ID NOS: 279
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 177
 ; LENGTH: 193
 ; TYPE: PRT
 ; ORGANISM: Homo sapien
 PCT-US03-38896-177

Query Match 1.0%; Score 8; DB 1; Length 193;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 693 KKKKKTNK 700
 DB 68 KKKKKTNK 75

RESULT 6
 US-09-978-360A-436
 ; Sequence 436, Application US/09978360A
 ; GENERAL INFORMATION:
 ; APPLICANT: Edwards, Jean-Baptiste Dumas Milne
 ; APPLICANT: Duclert, Aymeric
 ; APPLICANT: Bougueteloret, Lydie
 ; APPLICANT: Jobert, Severin

APPLICANT: Clusel, Catherine
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: 56, US4, CIP
; CURRENT APPLICATION NUMBER: US/09/978,360A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: US 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: US 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: US 60/099,273
; PRIOR FILING DATE: -09-04
; PRIOR APPLICATION NUMBER: US 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: US 09/215,435
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: PCT/1998/02122
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: US 09/247,155
; PRIOR FILING DATE: 1999-02-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 810
; SOFTWARE: Patent.pm
; SEQ ID NO 436
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -28...-1
; US-09-978-360A-436

Query Match 1.0%; Score 8; DB 5; Length 221;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LLAGVLLG 19
|||
Db 104 LLAGVLLG 111

RESULT 7
US-10-679-063-6997
; Sequence 6997, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
; CURRENT APPLICATION NUMBER: US/10/679,063
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: 60/415,758
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 27373
; SEQ ID NO 6997
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-679-063-6997

Query Match 1.0%; Score 8; DB 6; Length 221;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LLAGVLLG 19
|||
Db 104 LLAGVLLG 111

RESULT 8
US-10-679-063-8684
; Sequence 8684, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
; CURRENT APPLICATION NUMBER: US/10/679,063
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: 60/415,758
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 27373
; SEQ ID NO 8684
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-679-063-8684

Query Match 1.0%; Score 8; DB 6; Length 221;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LLAGVLLG 19
|||
Db 104 LLAGVLLG 111

RESULT 9
US-10-679-063-11202
; Sequence 11202, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
; CURRENT APPLICATION NUMBER: US/10/679,063
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: 60/415,758
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 27373
; SEQ ID NO 11202
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Papio anubis
US-10-679-063-11202

Query Match 1.0%; Score 8; DB 6; Length 221;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LLAGVLLG 19
|||
Db 104 LLAGVLLG 111

RESULT 10
US-10-425-114A-50595
; Sequence 50595, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 50595
; LENGTH: 222
; TYPE: PRT

ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: 700051071_FLI.pep
US-10-425-114A-50595

Query Match 1.0%; Score 8; DB 6; Length 222;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGARR 22
DB 210 VLLGARR 217

RESULT 11
US-10-425-114A-71744
Sequence 71744, Application US/10425114A
GENERAL INFORMATION:
APPLICANT: Liu, Jindong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114A
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 71744
LENGTH: 276
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: UC-ZMFLM017122B10_FLI.pep
US-10-425-114A-71744

Query Match 1.0%; Score 8; DB 6; Length 276;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 773 PRSPPTIC 780
DB 74 PRSPPTIC 81

RESULT 12
US-10-739-930-8397
Sequence 8397, Application US/10739930
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
FILE REFERENCE: 38-21(53377)B
CURRENT APPLICATION NUMBER: US/10/739,930
CURRENT FILING DATE: 2003-12-18
NUMBER OF SEQ ID NOS: 11088
SEQ ID NO 8397
LENGTH: 348
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: ZEAMA-23APR03-C6211_2.p
US-10-739-930-8397

Query Match 1.0%; Score 8; DB 6; Length 348;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LUGAARLP 24
DB 197 LUGAARLP 204

RESULT 13
PCT-US03-19153-296
Sequence 296, Application PC/TUS0319153
GENERAL INFORMATION:
APPLICANT: Diversa Corporation
TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM AND
TITLE OF INVENTION: METHODS FOR MAKING AND USING THEM
FILE REFERENCE: 09010-290W01
CURRENT APPLICATION NUMBER: PCT/US03/19153
CURRENT FILING DATE: 2003-06-16
PRIOR APPLICATION NUMBER: US 60/389,299
PRIOR FILING DATE: 2002-06-14
NUMBER OF SEQ ID NOS: 378
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 296
LENGTH: 377
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Obtained from an environmental sample.
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)...(26)
PCT-US03-19153-296

Query Match 1.0%; Score 8; DB 1; Length 377;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 FEGBSAT 234
DB 345 FEGBSAT 352

RESULT 14
US-10-463-720-296
Sequence 296, Application US/10463720
GENERAL INFORMATION:
APPLICANT: Steer, Brian
APPLICANT: Callen, Walter
APPLICANT: Healey, Shaun
APPLICANT: Hazlewood, Geoff
APPLICANT: Wu, Di
APPLICANT: Blum, David
APPLICANT: Esteghlalian, Alireza
TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM AND
TITLE OF INVENTION: METHODS FOR MAKING AND USING THEM
FILE REFERENCE: 09010-290W01
CURRENT APPLICATION NUMBER: US/10/463,720
CURRENT FILING DATE: 2003-06-16
PRIOR APPLICATION NUMBER: US 60/389,299
PRIOR FILING DATE: 2002-06-14
NUMBER OF SEQ ID NOS: 378
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 296
LENGTH: 377
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Obtained from an environmental sample.
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)...(26)
US-10-463-720-296

Query Match 1.0%; Score 8; DB 6; Length 377;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 FEGBSAT 234
DB 345 FEGBSAT 352

Db 345 PEGEGSAT 352

RESULT 15

US-10-425-114A-65537

; Sequence 65537, Application US/10425114A
; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(5313)B

; CURRENT APPLICATION NUMBER: US/10/425,114A

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 65537

; LENGTH: 485

; TYPE: PRT

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB4767-014-B10_FLI.pep

US-10-425-114A-65537

Query Match 1.0%; Score 8; DB 6; Length 485;

Best Local Similarity 100.0%; Pred.No. 78;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 653 LIFSCTLT 660

Db 107 LIFSCTLT 114

Search completed: February 20, 2004, 19:59:56.
Job time : 46 secs

This Page Blank (uspio)